

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 7, 2005, 10:58:39 ; Search time 1034.21 Seconds

(without alignments)

13.71.762 Million cell updates/sec

Title: US-09-786-502A-8

Perfect score: 30

Sequence: 1 tcgaggatcttgtcaaggaggatggctgc 30

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters:

9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_dat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description	FEATURES
1	30	100.0	30	6	BD272285		BD272285 Fused rec	Source
2	21	70.0	10951	1	AE005756		AB005756 Caulobact	
C	3	20.4	68.0	668	9	AF344837	AF344837 Pario cyn	
C	4	20.4	68.0	668	9	AF344842	AF344842 Cercobabu	
C	5	20.4	68.0	668	9	AF344845	AF344845 Callithri	
C	6	20.4	68.0	668	9	AF344852	AF344852 Macaca ne	
C	7	20.4	68.0	668	9	AF344855	AF344855 Macaca mu	
C	8	20.4	68.0	875	6	AR102241	AR102241 Sequence	
C	9	20.4	68.0	1514	6	BD267803	BD267803 Methods a	
C	10	20.4	68.0	1514	6	107254	107254 Sequence 11	
C	11	20.4	68.0	1514	6	AR380771	AR380771 Sequence	
C	12	20.4	68.0	1514	6	BD015187	BD015187 mSiSe	
C	13	20.4	68.0	1514	9	J02988	J02988 Human T-cell	
C	14	20.4	68.0	2405	6	AR102243	AR102243 Sequence	
C	15	20.4	68.0	3230	9	HUMCD284	M37815 Human T-cell	
C	16	20.4	68.0	3585	6	AR102242	AR102242 Sequence	
C	17	20.4	68.0	3803	6	AX557214	AX557214 Sequence	
C	18	20.4	68.0	3804	6	AX695882	AX695882 Sequence	
C	19	20.4	68.0	3806	6	CQ715869	CQ715869 Sequence	

RESULT 1		BD272285	LOCUS	DEFINITION	BD272285 Fused receptor specific to prostate-specific membrane antigen and its use.	
ACCESSION	BD272285	BD272285	VERSION	BD272285_1 GI:33082053	BD272285_1	JP 2002524081/A-8
KEYWORDS			SOURCE	synthetic construct	OS	Synthetic construct
			ORGANISM	synthetic construct	COMMENT	COMMENT
REFERENCE	1 (bases 1 to 30)	1 (bases 1 to 30)	AUTHORS	Sadelain,M., Bande,N.H. and Gong,M.	TITLE	TITLE
				Fused receptor specific to prostate-specific membrane antigen and its use.		
JOURNAL	Patent: JP 2002524081-A-8 06-AUG-2002;		JOURNAL	SLOAN KETTERING INSTITUTE FOR CANCER RESEARCH		
COMMENT	OS		COMMENT	OS		
PD	06-AUG-2002		PD	JP 2002524081-A/8		
PP	03-SEP-1999		PP	JP 20000568998		
PR	04-SEP-1998		PR	JP 20000568998		
PI	MICHEL SADELAIN NEIL H BANDER MICHAEL GONG		PI	60/099138		
PC	CD12N15/09, A61K35/26, A61K35/76, A61K39/395, A61K48/00, A61P35/00,		PC	PC07K19/00,		
PC			PC	PC12N5/10//C07K16/44, C12N15/00, C12N5/00		
CC			CC	Fused receptor specific to prostate-specific membrane antigen and its use		
CC			CC		Location/Qualifiers	Location/Qualifiers
CC			CC		1 .. 30	/organism='Synthetic construct'.
FT			FT			
FT			FT			

```
Query Match 100.0%; Score 30; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1	TCGAGGATCTTGTCAAGGGATAGGCCTC	30		CDS
Db	1	TCGAGGATCTTGTCAAGGGATAGGCCTC	30		
RESULT	2				
LOCUS	AE005756	Ae005756	10951 bp	DNA	linear BCT 12-JUN-2002
DEFINITION	Caulobacter crescentus	CB15	section 82 of	359 of the complete	
SENOMON					
ACCESSION	AE005756	AE005756			
VERSION	AB005756.1	GI : 13422042			
SOURCE	Caulobacter crescentus	CB15			
ORGANISM	Caulobacter crescentus	CB15			
BACTERIA; Proteobacteria; Alphaproteobacteria; Caulobacterales;					
REFERENCE	1 (bases 1 to 10951)				
AUTHORS	Nierman, W.C., Feldblyum, T.V., Laub, M.T., Paulsen, I.T., Nelson, K.E., Eisen, J., Potocka, I., Nelson, W.C., Newton, A., Alley, M.R., Ohta, N., Maddock, J.R., Phadke, N.D., Elly, B., DeBoy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M.B., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, C., Salzberg, S.L., Venter, J.C., Shapiro, L., and Fraser, C.M.				
TITLE	Complete genome sequence of Caulobacter crescentus				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)				
NEEDLINE	21173698				
PUBMED	11259647				
REFERENCE	2 (bases 1 to 10951)				
AUTHORS	Nierman, W.C., Feldblyum, T.V., Paulsen, I.T., Nelson, K.E., Eisen, J., Heidelberg, J.P., Alley, M.R.K., Ohta, N., Maddock, J.R., Potocka, I., Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D., Elly, B., Laub, M.T., DeBoy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M.B., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, C., Salzberg, S.L., Shapiro, L., Venter, J.C., and Fraser, C.M.				
JOURNAL	Submitted (31-JAN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
FEATURES	Location/Qualifiers				
1..10951	/organism="Caulobacter crescentus CB15"				
1	/mol_type="genomic DNA"				
1	/strain="CB15"				
1	/db_xref="Taxon:190650"				
177..2102	/gene="CC0798"				
177..2102	/note="Similar to GB:U40408 PID:1371055 PID:1628565				
1	/codon_start=1				
1	/transl_table=11				
1	/product="stalic acid-specific 9-O-acetyl esterase, putative"				
1	/protein_id="AAK22783..1"				
1	/db_xref="GI : 13422043"				
1	/translation="MSKKALITATAMGLSTPALAGPLASVSDHAYLQRDOPIRWGR				
1	APPNASTAVELSGQSGNBYPROALNGDGEVASAGDPDIRLQLTGKISRERPODSLPAQYWK				
1	GVWLGCSGSQNSBQYDGRVSAQDGRVSAVLPARPGGALTATARAGDETQVSNLVM				
1	VESTDQIANSFQACQHGRILKQDQGSGVSIQDWTSGLAALKYID				
1	GSLVLSYARDALGYARWSQSLGRVDAKLPAAKGTLQKQDDEAWKMPTEGWEQ				
1	APPDLAEGFDGTWLRLELTQKTAQAOQATLAQGPVDIDITFLNGREICSTPQYRDPTMSWQ				
1	RIVYLAQPLRKRGRNIALRKTIPQDQGSGVSIQDWTSGLAALKYID				
1	APLSESETALPPSPAPWLGASGLSLTRNGKTAIPVPGTKFAMYQGEANVTAQAYKRLM				
1	PRMIADORAFAFGGNOPFLLQTLAFCPEVDRPVNSAWAALRDYOROVALADPMVGM				
1	SPNDVGSFDIFPDAKRQGKQALQARKLAKGEAVASGAPLSARRGETTVVTL				
1	QAVVHRGNRGPQFELCDAACRFVDAVEGSSVRLATPAGMTATKVRAYADSPIN				
1	LFGANRURPATPPELVIP"				
1	complement (2112..3659)				
Db	1	TCGAGGATCTTGTCAAGGGATAGGCCTC	30		
RESULT	2				
LOCUS	AE005756	Ae005756	10951 bp	DNA	linear BCT 12-JUN-2002
DEFINITION	Caulobacter crescentus	CB15	section 82 of	359 of the complete	
SENOMON					
ACCESSION	AE005756	AE005756			
VERSION	AB005756.1	GI : 13422042			
SOURCE	Caulobacter crescentus	CB15			
ORGANISM	Caulobacter crescentus	CB15			
BACTERIA; Proteobacteria; Alphaproteobacteria; Caulobacterales;					
REFERENCE	1 (bases 1 to 10951)				
AUTHORS	Eisen, J., Heidelberg, J.P., Laub, M.T., Paulsen, I.T., Nelson, K.E., Potocka, I., Nelson, W.C., Newton, A., Alley, M.R., Ohta, N., Maddock, J.R., Phadke, N.D., Elly, B., DeBoy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M.B., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, C., Salzberg, S.L., Venter, J.C., Shapiro, L., and Fraser, C.M.				
TITLE	Complete genome sequence of Caulobacter crescentus				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)				
NEEDLINE	21173698				
PUBMED	11259647				
REFERENCE	2 (bases 1 to 10951)				
AUTHORS	Nierman, W.C., Feldblyum, T.V., Paulsen, I.T., Nelson, K.E., Eisen, J., Heidelberg, J.P., Alley, M.R.K., Ohta, N., Maddock, J.R., Potocka, I., Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D., Elly, B., Laub, M.T., DeBoy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M.B., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, C., Salzberg, S.L., Shapiro, L., Venter, J.C., and Fraser, C.M.				
JOURNAL	Submitted (31-JAN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
FEATURES	Location/Qualifiers				
1..10951	/organism="Caulobacter crescentus CB15"				
1	/mol_type="genomic DNA"				
1	/strain="CB15"				
1	/db_xref="Taxon:190650"				
177..2102	/gene="CC0798"				
177..2102	/note="Similar to GB:U40408 PID:1371055 PID:1628565				
1	/codon_start=1				
1	/transl_table=11				
1	/product="stalic acid-specific 9-O-acetyl esterase, putative"				
1	/protein_id="AAK22783..1"				
1	/db_xref="GI : 13422043"				
1	/translation="MSKKALITATAMGLSTPALAGPLASVSDHAYLQRDOPIRWGR				
1	APPNASTAVELSGQSGNBYPROALNGDGEVASAGDPDIRLQLTGKISRERPODSLPAQYWK				
1	GVWLGCSGSQNSBQYDGRVSAQDGRVSAVLPARPGGALTATARAGDETQVSNLVM				
1	VESTDQIANSFQACQHGRILKQDQGSGVSIQDWTSGLAALKYID				
1	GSLVLSYARDALGYARWSQSLGRVDAKLPAAKGTLQKQDDEAWKMPTEGWEQ				
1	APPDLAEGFDGTWLRLELTQKTAQAOQATLAQGPVDIDITFLNGREICSTPQYRDPTMSWQ				
1	RIVYLAQPLRKRGRNIALRKTIPQDQGSGVSIQDWTSGLAALKYID				
1	APLSESETALPPSPAPWLGASGLSLTRNGKTAIPVPGTKFAMYQGEANVTAQAYKRLM				
1	PRMIADORAFAFGGNOPFLLQTLAFCPEVDRPVNSAWAALRDYOROVALADPMVGM				
1	SPNDVGSFDIFPDAKRQGKQALQARKLAKGEAVASGAPLSARRGETTVVTL				
1	QAVVHRGNRGPQFELCDAACRFVDAVEGSSVRLATPAGMTATKVRAYADSPIN				
1	LFGANRURPATPPELVIP"				
1	complement (2112..3659)				
Db	1	TCGAGGATCTTGTCAAGGGATAGGCCTC	30		
RESULT	2				
LOCUS	AE005756	Ae005756	10951 bp	DNA	linear BCT 12-JUN-2002
DEFINITION	Caulobacter crescentus	CB15	section 82 of	359 of the complete	
SENOMON					
ACCESSION	AE005756	AE005756			
VERSION	AB005756.1	GI : 13422042			
SOURCE	Caulobacter crescentus	CB15			
ORGANISM	Caulobacter crescentus	CB15			
BACTERIA; Proteobacteria; Alphaproteobacteria; Caulobacterales;					
REFERENCE	1 (bases 1 to 10951)				
AUTHORS	Eisen, J., Heidelberg, J.P., Laub, M.T., Paulsen, I.T., Nelson, K.E., Potocka, I., Nelson, W.C., Newton, A., Alley, M.R., Ohta, N., Maddock, J.R., Phadke, N.D., Elly, B., DeBoy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M.B., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, C., Salzberg, S.L., Venter, J.C., Shapiro, L., and Fraser, C.M.				
TITLE	Complete genome sequence of Caulobacter crescentus				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)				
NEEDLINE	21173698				
PUBMED	11259647				
REFERENCE	2 (bases 1 to 10951)				
AUTHORS	Nierman, W.C., Feldblyum, T.V., Paulsen, I.T., Nelson, K.E., Eisen, J., Heidelberg, J.P., Alley, M.R.K., Ohta, N., Maddock, J.R., Potocka, I., Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D., Elly, B., Laub, M.T., DeBoy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M.B., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, C., Salzberg, S.L., Shapiro, L., Venter, J.C., and Fraser, C.M.				
JOURNAL	Submitted (31-JAN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
FEATURES	Location/Qualifiers				
1..10951	/organism="Caulobacter crescentus CB15"				
1	/mol_type="genomic DNA"				
1	/strain="CB15"				
1	/db_xref="Taxon:190650"				
177..2102	/gene="CC0798"				
177..2102	/note="Similar to GB:U40408 PID:1371055 PID:1628565				
1	/codon_start=1				
1	/transl_table=11				
1	/product="stalic acid-specific 9-O-acetyl esterase, putative"				
1	/protein_id="AAK22783..1"				
1	/db_xref="GI : 13422043"				
1	/translation="MSKKALITATAMGLSTPALAGPLASVSDHAYLQRDOPIRWGR				
1	APPNASTAVELSGQSGNBYPROALNGDGEVASAGDPDIRLQLTGKISRERPODSLPAQYWK				
1	GVWLGCSGSQNSBQYDGRVSAQDGRVSAVLPARPGGALTATARAGDETQVSNLVM				
1	VESTDQIANSFQACQHGRILKQDQGSGVSIQDWTSGLAALKYID				
1	GSLVLSYARDALGYARWSQSLGRVDAKLPAAKGTLQKQDDEAWKMPTEGWEQ				
1	APPDLAEGFDGTWLRLELTQKTAQAOQATLAQGPVDIDITFLNGREICSTPQYRDPTMSWQ				
1	RIVYLAQPLRKRGRNIALRKTIPQDQGSGVSIQDWTSGLAALKYID				
1	APLSESETALPPSPAPWLGASGLSLTRNGKTAIPVPGTKFAMYQGEANVTAQAYKRLM				
1	PRMIADORAFAFGGNOPFLLQTLAFCPEVDRPVNSAWAALRDYOROVALADPMVGM				
1	SPNDVGSFDIFPDAKRQGKQALQARKLAKGEAVASGAPLSARRGETTVVTL				
1	QAVVHRGNRGPQFELCDAACRFVDAVEGSSVRLATPAGMTATKVRAYADSPIN				
1	LFGANRURPATPPELVIP"				
1	complement (2112..3659)				
Db	1	TCGAGGATCTTGTCAAGGGATAGGCCTC	30		
RESULT	2				
LOCUS	AE005756	Ae005756	10951 bp	DNA	linear BCT 12-JUN-2002
DEFINITION	Caulobacter crescentus	CB15	section 82 of	359 of the complete	
SENOMON					
ACCESSION	AE005756	AE005756			
VERSION	AB005756.1	GI : 13422042			
SOURCE	Caulobacter crescentus	CB15			
ORGANISM	Caulobacter crescentus	CB15			
BACTERIA; Proteobacteria; Alphaproteobacteria; Caulobacterales;					
REFERENCE	1 (bases 1 to 10951)				
AUTHORS	Eisen, J., Heidelberg, J.P., Laub, M.T., Paulsen, I.T., Nelson, K.E., Potocka, I., Nelson, W.C., Newton, A., Alley, M.R., Ohta, N., Maddock, J.R., Phadke, N.D., Elly, B., DeBoy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M.B., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, C., Salzberg, S.L., Venter, J.C., Shapiro, L., and Fraser, C.M.				
TITLE	Complete genome sequence of Caulobacter crescentus				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (7), 4136-4141 (2001)				
NEEDLINE	21173698				
PUBMED	11259647				
REFERENCE	2 (bases 1 to 10951)				
AUTHORS	Nierman, W.C., Feldblyum, T.V., Paulsen, I.T., Nelson, K.E., Eisen, J., Heidelberg, J.P., Alley, M.R.K., Ohta, N., Maddock, J.R., Potocka, I., Nelson, W.C., Newton, A., Stephens, C., Phadke, N.D., Elly, B., Laub, M.T., DeBoy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M.B., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, C., Salzberg, S.L., Shapiro, L., Venter, J.C., and Fraser, C.M.				
JOURNAL	Submitted (31-JAN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
FEATURES	Location/Qualifiers				
1..10951	/organism="Caulobacter crescentus CB15"				
1	/mol_type="genomic DNA"				
1	/strain="CB15"				
1	/db_xref="Taxon:190650"				
177..2102	/gene="CC0798"				
177..2102	/note="Similar to GB:U40408 PID:1371055 PID:1628565				
1	/codon_start=1				
1	/transl_table=11				
1	/product="stalic acid-specific 9-O-acetyl esterase, putative"				
1	/protein_id="AAK22783..1"				
1	/db_xref="GI : 13422043"				
1	/translation="MSKKALITATAMGLSTPALAGPLASVSDHAYLQRDOPIRWGR				
1	APPNASTAVELSGQSGNBYPROALNGDGEVASAGDPDIRLQLTGKISRERPODSLPAQYWK				
1	GVWLGCSGSQNSBQYDGRVSAQDGRVSAVLPARPGGALTATARAGDETQVSNLVM				
1	VESTDQIANSFQACQHGRILKQDQGSGVSIQDWTSGLAALKYID				
1	GSLVLSYARDALGYARWSQSLGRVDAKLPAAKGTLQKQDDEAWKMPTEGWEQ				
1	APPDLAEGFDGTWLRLELTQKTAQAOQATLAQGPVDIDITFLNGREICSTPQYRDPTMSWQ				
1	RIVYLAQPLRKRGRNIALRKTIPQDQGSGVSIQDWTSGLAALKYID				
1	APLSESETALPPSPAPWLGASGLSLTRNGKTAIPVPGTKFAMYQGEANVTAQAYKRLM				
1	PRMIADORAFAFGGNOPFLLQTLAFCPEVDRPVNSAWAALRDYOROVALADPMVGM				
1	SPNDVGSFDIFPDAKRQGKQALQARKLAKGEAVASGAPLSARRGETTVVTL				
1	QAVVHRGNRGPQFELCDAACRFVDAVEGSSVRLATPAGMTATKVRAYADSPIN				
1	LFGANRURPATPPELVIP"				
1	complement (2112..3659)				
Db	1	TCGAGGATCTTGTCAAGGGATAGGCCTC	30		
RESULT	2				
LOCUS	AE005756	Ae005756	10951 bp	DNA	linear BCT 12-JUN-2002
DEFINITION	Caulobacter crescentus	CB15	section 82 of	359 of the complete	
SENOMON					
ACCESSION	AE005756	AE005756			
VERSION	AB005756.1	GI : 13422042			
SOURCE	Caulobacter crescentus	CB15			
ORGANISM	Caulobacter crescentus	CB15			
BACTERIA; Proteobacteria; Alphaproteobacteria; Caulobacterales;					
REFERENCE	1 (bases 1 to 10951)				
AUTHORS	Eisen, J., Heidelberg, J.P., Laub, M.T., Paulsen, I.T., Nelson, K.E., Potocka, I., Nelson, W.C., Newton, A., Alley, M.R., Ohta, N., Maddock, J.R., Phadke, N.D., Elly, B., DeBoy, R.T., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Kolonay, J.F., Smit, J., Craven, M.B., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, C., Salzberg, S.L., V				

/gene="CC0803" /note="identified by Glimmer2; putative" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAK22788_1" /db_xref="GI:13422048" /translation="MADGEFDHIDQEQAARIKKVADRLGMSVSEYAVALIDAIGLGAAPRAIDPDAIDAIADIERGEPATSRDEFAHIRVTAGLG"	AUTHORS Villinger,F. TITLE Direct Submission JOURNAL Submitted (02-FEB-2001) Pathology and Laboratory Medicine, Emory University School of Medicine, Winship Cancer Institute, Clifton Rd, Atlanta, GA 30322, USA FEATURES Location/Qualifiers 1. .668 Source /organism="Papio anubis" /mol_type="mRNA" /sub_species="anubis" /db_xref="taxon:9555" 2. .664 /note="co-stimulatory molecule" /codon_start=1 /product="CD28 protein precursor" /protein_id="AAK3753_1" /db_xref="GI:13649991" /translation="MLRILALAINLNLPISQIYTGNKILVKQSPMLVAYDNVNLSCKSYVNLFSRERASDHKGLDSSAVECVVVGNYPPYLDENSNGTIIHVKGHLCPPLFPPPSKPPWALVYVGYNODDIYFCKLEVNYPYPPYLDENSNGTIIHVKGHLCPPLFPPPSKPPWALVYVGVLACYSLLVIVAFSIFCMRSKRSRLLHSYMMTPRRGPTRKHYQYAPPDFAYRS"
gene CDS	ORIGIN 2. .55
complement ("610. .10699")	Query Match 68.0%; Score 20.4; DB 9; Best Local Similarity 95.5%; Pred. No. 1.1e+02; Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
complement ("610. .10699")	Qy 9 CTTGTCAAGGCCATAGGCTGC 30 Db 668 CGTGTCAAGGCCATAGGCTGC 647
dehydrogenase	RESULT 4 AF344842/C LOCUS AF344842 DEFINITION Cercocebus torquatus atys CD28 protein mRNA, complete cds. ACCESSION AF344842 VERSION AF344842.1 KEYWORDS Gaps/Fas-ligand and co-stimulatory molecules SOURCE Cercocebus torquatus atys (sooty mangabey) ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; REFERENCE 1 (bases 1 to 668) AUTHORS Villinger,F., Bostik,P., Mayne,A., King,C.L., Genain,C.P., Weiss,W.R. and Ansari,A.A. TITLE Cloning, sequencing, and homology analysis of nonhuman primate Fas/Fas-ligand and co-stimulatory molecules JOURNAL Immunogenetics 53 (4), 315-328 (2001) MEDLINE 21383618 PUBMED 11491535 REPERB 2 (bases 1 to 668)
dehydrogenase	FEATURES Location/Qualifiers 1. .668 /organism="Cercocebus torquatus atys" /mol_type="mRNA" /sub_species="atys" /db_xref="taxon:9531" 2. .664 /codon_start=1 /product="CD28 protein" /protein_id="AAK3760_1" /db_xref="GI:13655471" /translation="MLRILALAINLNLPISQIYTGNKILVKQSPMLVAYDNVNLSCKSYVNLFSRERASDHKGLDSSAVECVVVGNYPPYLDENSNGTIIHVKGHLCPPLFPPPSKPPWALVYVGVLACYSLLVIVAFSIFCMRSKRSRLLHSYMMTPRRGPTRKHYQYAPPDFAYRS"
protein_id="AAK22789_1" /note="identified by match to PFAM protein family HMM PRO1619" /db_xref="GI:13422049" /product="proline/depta-1-pyrroline-5-carboxylate dehydrogenase" /note="dehydrogenase" /protein_id="AAK22789_1"	ORIGIN Query Match 70.0%; Score 21; DB 1; Length 10951; Best Local Similarity 82.8%; Pred. No. 75; Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
protein_id="AAK22789_1"	Qy 1 TCGAGGATCTTGTCAGGAGCGATAGGCTG 29 Db 52299 TCAGGATCTGTCAGGATGCTGGCTG 5257
PRO1619	RESULT 3 AF344837/C LOCUS AF344837 DEFINITION Papio cynocephalus anubis CD28 protein precursor, mRNA, complete cds. ACCESSION AF344837 VERSION AF344837.1 SOURCE Papio anubis (olive baboon) ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Papio. REFERENCE 1 (bases 1 to 668) AUTHORS Villinger,F., Bostik,P., Mayne,A., King,C.L., Genain,C.P., Weiss,W.R. and Ansari,A.A. TITLE Cloning, sequencing, and homology analysis of nonhuman primate Fas/Fas-ligand and co-stimulatory molecules JOURNAL Immunogenetics 53 (4), 315-328 (2001) MEDLINE 21383618 PUBMED 11491535 REFERENCE 2 (bases 1 to 668)

ORIGIN	YVQTDIYFCKIEVMPPYLDNEKSNTIIHTVKHLCPSPFPGPSKPFWALVVG GVLAICYSLVTVAFRIFWMSRSRLHSDYNNMTPRRPGPTRKHQFYAPPDFAA RS"	PRI 06-SEP-2001	AP244852.1	GI:13655492
VERSION	Macaca nemestrina (pig-tailed macaque)			
KEYWORDS				
SOURCE				
ORGANISM	Macaca nemestrina; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Cercopithecinae; Macaca.			
REFERENCE	1 (bases 1 to 668)			
AUTHORS	Villinge, F., Bostik, P., Mayne, A., King, C.L., Genain, C.P., Weiss, W.R. and Ansari, A.A.			
TITLE	Cloning, sequencing, and homology analysis of nonhuman primate Fas/Fas-ligand and co-stimulatory molecules			
JOURNAL	Immunogenetics 53 (4), 315-328 (2001)			
MEDLINE	21383618			
PUBMED	11491535			
REFERENCE	2 (bases 1 to 668)			
AUTHORS	Villinge, F.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-FEB-2001) Pathology and Laboratory Medicine, Emory University School of Medicine, Winship Cancer Institute, 1365B Clifton Rd, Atlanta, GA 30322, USA			
FEATURES				
source	Location/Qualifiers			
1.	668			
/organism="Macaca nemestrina"				
/mol_type="mRNA"				
/db_xref="taxon:9545"				
CDS				
1.	664			
/note="co-stimulatory molecule"				
/product="CD28 protein precursor"				
/protein_id="AAK37612.1"				
/codon_start=1				
/db_xref="IGI:13655493"				
/translation=MURLIAILNLPSIQVTGNKLIVVKQSPLMLVAYDNVNLSCKSY YVNQDIFYECKIEVMPYPLDNERSNGTIVHVKGKHLCPSPLEPSPSKPWPWLVNG GVLACYSLLTVAPCIFWMSRSRLHSDYNNMTPRRPGPTRKHQFYAPPDFAA RS"				
sig_peptide	2 . .55			
variation	503			
variation	645			
ORIGIN				
Query Match	68.0%	Score 20.4;	DB 9;	Length 668;
Best Local Similarity	95.5%	Pred. No. 1.1e+02;		
Matches	21;	Conservative	0;	Indels 0; Gaps 0;
DEFINITION	Macaca mulatta CD28 protein precursor, mRNA, complete cds.			
VERSION	AF244855			
KEYWORDS	Macaca mulatta (rhesus monkey)			
SOURCE				
ORGANISM	Macaca mulatta; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Cercopithecinae; Macaca.			
REFERENCE	1 (bases 1 to 668)			
AUTHORS	Villinge, F., Bostik, P., Mayne, A., King, C.L., Genain, C.P., Weiss, W.R. and Ansari, A.A.			
TITLE	Cloning, sequencing, and homology analysis of nonhuman primate Fas/Fas-ligand and co-stimulatory molecules			
JOURNAL	Immunogenetics 53 (4), 315-328 (2001)			
MEDLINE	21383618			
PUBMED	11491535			
REFERENCE	2 (bases 1 to 668)			
AUTHORS	Villinge, F.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-FEB-2001) Pathology and Laboratory Medicine, Emory University School of Medicine, Winship Cancer Institute, 1365B Clifton Rd, Atlanta, GA 30322, USA			
FEATURES				
source	Location/Qualifiers			
1.	668			
/organism="Callithrix jacchus"				
/mol_type="mRNA"				
/db_xref="taxon:9483"				
CDS				
2.	664			
/note="co-stimulatory molecule"				
/product="CD28 protein"				
/protein_id="AAK37604.1"				
/codon_start=1				
/db_xref="IGI:13655477"				
/translation=MURLLIAILNLPSIQVTGNKLIVVKQSPLMLVAYDNVNLSCKSY NLFSPLOATGKILYKQSPMLAYDNTVNLTCKSY YVNQDIFYCCKIEVMPYPLDNERSNGTIVHVKGKHLCPSPLEPSPSKPWPWLVNG GVLACYSLLTVAPCIFWMSRSRLHSDYNNMTPRRPGPTRKHQFYAPPDFAA RS"				
sig_peptide	2 . .55			
ORIGIN				
Query Match	68.0%	Score 20.4;	DB 9;	Length 668;
Best Local Similarity	95.5%	Pred. No. 1.1e+02;		
Matches	21;	Conservative	0;	Indels 0; Gaps 0;
DEFINITION	Macaca mulatta (rhesus monkey)			
VERSION	AF244855			
KEYWORDS	Macaca mulatta CD28 protein precursor, mRNA, linear			
SOURCE				
ORGANISM	Macaca mulatta; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Cercopithecinae; Macaca.			
REFERENCE	1 (bases 1 to 668)			
AUTHORS	Villinge, F., Bostik, P., Mayne, A., King, C.L., Genain, C.P., Weiss, W.R. and Ansari, A.A.			
TITLE	Cloning, sequencing, and homology analysis of nonhuman primate Fas/Fas-ligand and co-stimulatory molecules			
JOURNAL	Immunogenetics 53 (4), 315-328 (2001)			
MEDLINE	21383618			
PUBMED	11491535			
RESULT	6			
LOCUS	AF344852/C			
DEFINITION	Macaca nemestrina CD28 protein Precursor			
ACCESSION	AF344852			
VERSION	AF344852			
KEYWORDS	Macaca nemestrina, complete cds.			
SOURCE				
ORGANISM	Macaca nemestrina (pig-tailed macaque)			
REFERENCE	1 (bases 1 to 668)			
AUTHORS	Villinge, F., Bostik, P., Mayne, A., King, C.L., Genain, C.P., Weiss, W.R. and Ansari, A.A.			
TITLE	Cloning, sequencing, and homology analysis of nonhuman primate Fas/Fas-ligand and co-stimulatory molecules			
JOURNAL	Immunogenetics 53 (4), 315-328 (2001)			
MEDLINE	21383618			
PUBMED	11491535			

REFERENCE 2 (bases 1 to 668)
 AUTHORS Villinger, F.
 TITLE Direct Submission
 JOURNAL Submitted (02-FEB-2001) Pathology and Laboratory Medicine, Emory University School of Medicine, Winship Cancer Institute, 1365B Clifton Rd, Atlanta, GA 30322, USA
 FEATURES source
 1..668
 /organism="Macaca mulatta"
 /mol_type="mRNA"
 /db_xref="taxon:9544"
 2..664
 /note="co-stimulatory molecule"
 /codon_start=1
 /product="C928 protein precursor"
 /protein_id="AAR37538_1"
 /db_xref="GI:13650008"
 /translation="MRLILALNLNLPSIOTNGNKLIVKQSPMLVAYDNAYNLSSCKSYVNLFSREFASLKLIGLSDAVEVVYEDYSQQQLVTSKTEFCGDKLGNESTVFLQNLVNTQDLYFCKLEKVNNPPYDNEKEKGNTTIEKGKHLCPSPLPFRPSDFWALVYVGWVLAISLLVTVTAFSIFWMRSRSRSLHSDYMMNTPERRPGTRKHYQPCAPPDFAYRS"
 sig_peptide 2..55
 variation 55
 /replace="t"
 variation 102
 /replace="g"
 variation 142
 /replace="t"
 variation 210
 /replace="g"
 variation 212
 /replace="a"
 variation 249
 /replace="g"
 variation 365
 /replace="a"
 variation 524
 /replace="t"
 variation 547
 /replace="g"
 variation 627
 /replace="a"
 /replace="a"
 ORIGIN
 Query Match 68.0%; Score 20.4; DB 9;
 Best Local Similarity 95.5%; Pred. No. 1.e+02;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 9 CTTGTCAAGGAGCGATAGGTGTC 30
 Db 668 CGTGTCAAGGAGCGATAGGTGTC 647

RESULT 8
 AR102241/_C LOCUS AR102241 875 bp DNA linear PAT 14-FEB-2001
 DEFINITION Sequence 1 from patent US 6083751.
 ACCESSION AR102241
 VERSION AR102241.1 GI:12813039
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 875)
 AUTHORS Feldhaus, A, Lawrence, and Jones L Ann.
 TITLE Chimeric receptors for the generation of selectively-activatable T cells
 JOURNAL TH-independent cytotoxic T cells
 Patent: US 6083751-A 1 04-JUL-2000;
 FEATURES source
 1..875
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 68.0%; Score 20.4; DB 6;
 Best Local Similarity 95.5%; Pred. No. 1.2e+02;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 9 CTTGTCAAGGAGCGATAGGTGTC 30
 Db 766 CGTGTCAAGGAGCGATAGGTGTC 745

RESULT 10
 I07254/_C LOCUS I07254 1514 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 11 from Patent EP 0320191.
 ACCESSION I07254
 VERSION I07254.1 GI:590064
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 1514)
 AUTHORS Seed, B.D., Do.M.B., Allen, J., Aruffo, A., Camerini, D., Lauffer, L.D., Oquendo, C.P., Simmons, D., Stamenkovic, I., and Stengelin, S.D.

Query Match 68.0%; Score 20.4; DB 9; Length 1514;
 Best Local Similarity 95.5%; Pred. No. 1.2e+02; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 1;

LOCUS AR102243/C 2405 bp DNA linear PAT 14-FEB-2001
 DEFINITION Sequence 3 from Patent US 6083751.
 ACCESSION AR102243
 VERSION AR102243.1 GI:12813041
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1. (bases 1 to 2405)
 AUTHORS Feldhaus,A,Lawrence, and Jones,L.Ann.
 TITLE Chimeric receptors for the generation of selectively-activatable TH-independent cytotoxic T cells
 JOURNAL Patent: US 6083751-A 3 04-JUL-2000;
 FEATURES Location/Qualifiers 1..2405
 SOURCE /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Query Match 68.0%; Score 20.4; DB 6; Length 2405;
 Best Local Similarity 95.5%; Pred. No. 1.2e+02; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 1;

LOCUS AR102243/C 2405 bp DNA linear PAT 14-FEB-2001
 DEFINITION Sequence 3 from Patent US 6083751.
 ACCESSION AR102243
 VERSION AR102243.1 GI:12813041
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unclassified.

RESULT 14 HUMCD28A/C 3230 bp DNA linear PRI 01-NOV-1994
 LOCUS HUMCD28A
 DEFINITION Human T-cell membrane glycoprotein CD28 mRNA, exon 4.
 ACCESSION M37815
 VERSION M37815.1 GI:1600089
 KEYWORDS CD28
 SEGMENT 4 of 4
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.
 REFERENCE Lee,K.P., Taylor,C., Petryniak,B., Turka,L.A., June,C.H. and Thompson,C.B.
 TITLE The genomic organization of the CD28 gene. Implications for the regulation of CD28 mRNA expression and heterogeneity
 JOURNAL J. Immunol. 145 (1), 344-352 (1990)
 MEDLINE 90293482
 PUBMED 2162832
 COMMENT Original source text: Human DNA.
 FEATURES Location/Qualifiers 1..3230
 SOURCE /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="Q33;Q34"
 /gene="CD28"
 /join(M37812..1..539..812..M37813..1..1..542..M37814..1..1..467,
 122..3169)

gene mRNA

Query Match 68.0%; Score 20.4; DB 9; Length 3230;
 Best Local Similarity 95.5%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 1;

LOCUS AR102242/C 3585 bp DNA linear PAT 14-FEB-2001
 DEFINITION Sequence 2 from patent US 6083751.
 ACCESSION AR102242
 VERSION AR102242.1 GI:12813040
 KEYWORDS Unknown.
 SOURCE Organism
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 3585)
 AUTHORS Pfeilhau, A.Lawrence, and Jones,L.Ann.
 TITLE Chimeric receptors for the generation of selectively-activatable TH-independent cytotoxic T cells
 JOURNAL Patent: US 6083751-A 2 04-JUL-2000;

FEATURES	Location/Qualifiers 1. .3585 /organism="unknown" /mol_type="unassigned DNA"							
ORIGIN								
Query Match	Score 20.4; DB 6; Best Local Similarity 95.5%; Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	Length 3585;			Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	9 CTTGTCAGGAGCGATAAGCTGC 30				Qy 9 CTTGTCAGGAGCGATAAGCTGC 30			
Db	3578 CGTGTCAAGCGATAGCTGC 3557				Db 889 CGTGTCAAGCGATAGCTGC 868			
RESULT 17								
AX557214/C	AX557214 LOCUS Sequence 1 from Patent WO02066059.	DNA linear	PAT 27-NOV-2002		RESULT 19 CQ715869/C LOCUS Sequence 1 from Patent WO02068579.	DNA linear	PAT 03-FEB-2004	
DEFINITION AX557214					DEFINITION Sequence 1803 from Patent WO02068579.			
ACCESSION AX557214.1	GI:25900195				ACCESSION CQ715869.1 GI:42276726			
VERSION					KEYWORDS			
SOURCE	Homo sapiens (human)				SOURCE Homo sapiens (human)			
ORGANISM	Homo sapiens				ORGANISM Homo sapiens			
	Mammalia; Eutheria; Primates; Chordata; Craniata; Vertebrata; Euteleostomi;				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE 1					REFERENCE 1			
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.					AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.			
TITLE					TITLE			
	such as nucleic acid arrays, comprising a majority of humanoxes or transcripts, for detecting expression and other uses thereof							
JOURNAL					JOURNAL Patent: WO 02068579-A 1803 06-SEP-2002;			
					PEPCorporation (NY) (US)			
FEATURES					FEATURES Location/Qualifiers			
SOURCE	1. .3806				SOURCE 1. .3806			
	/mol type="unassigned DNA"					/mol type="unassigned DNA"		
ORIGIN					ORIGIN /db_xref="taxon:9606"			
Query Match	Score 68.0%;	Score 20.4;	DB 6;	Length 3806;				
Best Local Similarity 95.5%;	Pred. No. 1.3e+02;							
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;								
Qy	9 CTTGTCAGGAGCGATAAGCTGC 30				Qy 9 CTTGTCAGGAGCGATAAGCTGC 30			
Db	889 CGTGTCAAGCGATAGCTGC 868				Db 889 CGTGTCAAGCGATAGCTGC 868			
RESULT 20					RESULT 20			
AX283559/C					AX283559/C			
DEFINITION Sequence 3 from Patent WO019300.					DEFINITION Sequence 3 from Patent WO019300.			
ACCESSION AX283559					ACCESSION AX283559			
VERSION					VERSION AX283559.1 GI:17044297			
SOURCE	Homo sapiens (human)				KEYWORDS			
ORGANISM	Homo sapiens				SOURCE Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				ORGANISM Homo sapiens			
REFERENCE 1					REFERENCE 1			
AUTHORS Bluestone, J.A., Collins, M., Whitters, M., Griffin, M. and Kranz, D.					AUTHORS Bluestone, J.A., Collins, M., Whitters, M., Griffin, M. and Kranz, D.			
TITLE Surface-bound antigen binding portions of antibodies that bind to ctla-4 and cd28 and uses therefor					TITLE			
JOURNAL Patent: WO 0179300-A 3 25-OCT-2001;					JOURNAL GENETICS INSTITUTE, INC. (US)			
					LOCATION/Qualifiers			
FEATURES	1. .3806				FEATURES 1. .3806			
SOURCE	/organism="Homo sapiens"					/organism="Homo sapiens"		
	/mol type="unassigned DNA"					/mol type="unassigned DNA"		
ORIGIN	/db_xref="taxon:9606"				ORIGIN /db_xref="taxon:9606"			
Query Match	Score 68.0%;	Score 20.4;	DB 6;	Length 3806;				
Best Local Similarity 95.5%;	Pred. No. 1.3e+02;							
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;								
Qy	9 CTTGTCAGGAGCGATAAGCTGC 30				Qy 9 CTTGTCAGGAGCGATAAGCTGC 30			
Db	889 CGTGTCAAGCGATAGCTGC 867				Db 889 CGTGTCAAGCGATAGCTGC 868			
RESULT 18					RESULT 18			
AX95882/C	AX95882 LOCUS Sequence 1509 from Patent WO03008583.	DNA linear	PAT 31-MAR-2003		AX95882/C LOCUS Sequence 3 from Patent WO019300.			
DEFINITION Sequence 1509 from Patent WO03008583.					DEFINITION Sequence 3 from Patent WO019300.			
ACCESSION AX95882					ACCESSION AX283559			
VERSION AX95882.1	GI:29419047				VERSION AX283559			
KEYWORDS					KEYWORDS			
SOURCE	Homo sapiens (human)				SOURCE Homo sapiens			
ORGANISM	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				ORGANISM Homo sapiens			
REFERENCE 1					REFERENCE 1			
AUTHORS Morris, D.W. and Engelhardt, E.K.					AUTHORS Morris, D.W. and Engelhardt, E.K.			
TITLE Novel compositions and methods for cancer					TITLE Novel compositions and methods for cancer			
JOURNAL Sagres Discovery (US)					JOURNAL Sagres Discovery (US)			
FEATURES	Location/Qualifiers				FEATURES Location/Qualifiers			
Source	1. .3804				Source 1. .3804			
	/organism="Homo sapiens"					/organism="Homo sapiens"		
	/mol type="unassigned DNA"					/mol type="unassigned DNA"		
ORIGIN	/db_xref="taxon:9606"				ORIGIN /db_xref="taxon:9606"			
Query Match	Score 68.0%;	Score 20.4;	DB 6;	Length 3804;				
Best Local Similarity 95.5%;	Pred. No. 1.3e+02;							
Qy	9 CTTGTCAGGAGCGATAAGCTGC 30				Qy 9 CTTGTCAGGAGCGATAAGCTGC 30			
Db	889 CGTGTCAAGCGATAGCTGC 868				Db 889 CGTGTCAAGCGATAGCTGC 868			

RESULT 21
 AX695881/c AX695881 51365 bp DNA linear PAT 31-MAR-2003
 DEFINITION Sequence 1508 from Patent WO2003008583.
 ACCESSION AX655881
 VERSION GI:29419046
 KEYWORDS Homo sapiens (human)
 SOURCE ORGANISM Homo sapiens
 Bukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Morris D.W. and Engelhard, E.K.
 TITLE Novel compositions and methods for cancer
 JOURNAL Patent: WO 0308583-A 1508 30-JAN-2003;
 Sagres Discovery (US)
 Location/Qualifiers 1..51365
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN Query Match_Similarity 68.0%; Score 20.4; DB 6; Length 51365;
 Best Local_Similarity 95.5%; Pred. No. 1..7e+02;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 9 CTTGTCAGGAGCGATAGCCCTGC 30
 Db 38448 CCTGTCAGGAGCGATAGCCCTGC 384227

RESULT 22
 AF225899/c AF225899 106539 bp DNA linear PRI 23-FEB-2000
 DEFINITION Homo sapiens PAC clone 219d7, complete sequence.
 VERSION AF225899
 KEYWORDS SOURCE ORGANISM Homo sapiens (human)
 Homo sapiens
 Bukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCES 1
 (bases 1 to 106539)
 AUTHORS Fitzpatrick, E., Hammond, H.A., DeAngelis, D.M., Soderman, A.R.,
 Wright, J.L., Liu, X., Larson, D., McGowan, J., Ziegler, S.,
 Pritchard, L., Hess, J.P., Todd, J., Caskey, C.T. and Metzker, M.L.
 TITLE Direct Submission
 JOURNAL Submitted (18-JAN-2000) Department of Human Genetics, Merck & Co.,
 Inc., SunneyTown Pike, West Point, PA 19486, USA
 FEATURES SOURCE
 1..106539
 /organism="Homo sapiens"
 /mol_type="Genomic DNA"
 /db_xref="taxon:9606"
 /clone="PAC_219d7"
 repeat_region complement(34..104)
 /rpt_family="MERS5"
 complement(384..662)
 /rpt_family="Alu"
 repeat_region 1021..1304
 /rpt_family="Alu"
 repeat_region 180..2135
 /rpt_family="Alu"
 repeat_region 2139..2439
 /rpt_family="Alu"
 repeat_region 3214..3294
 /rpt_family="MLT2B2"
 repeat_region 3588..3657
 /rpt_family="MLT2B2"
 repeat_region 3982..4274
 /rpt_family="Alu"
 repeat_region /rpt_family="Alu"
 repeat_region complement(5956..59701)

repeat_region 5406..5550
 /rpt_family="MERS3"
 5441..5509
 /rpt_family="MERS3"
 complement(5558..5834)
 /rpt_family="Alu"
 repeat_region 5851..5998
 /rpt_family="MERS3"
 complement(6534..6827)
 /rpt_family="Alu"
 complement(8239..8568)
 /rpt_family="STR12"
 repeat_region 11884..12159
 /rpt_family="Alu"
 complement(12838..13137)
 /rpt_family="Alu"
 complement(14450..14691)
 /rpt_family="Alu"
 complement(15004..15291)
 /rpt_family="Alu"
 complement(16143..16422)
 /rpt_family="Alu"
 complement(16513..17096)
 /rpt_family="Alu"
 complement(1752..17706)
 /rpt_family="MIR"
 repeat_region 21610..21890
 /rpt_family="Alu"
 complement(2793..28279)
 /rpt_family="Alu"
 repeat_region 28842..28933
 /rpt_family="MIR"
 STS 30459..23329
 /note="Li8220 Human chromosome 2 STS UT426"
 repeat_region 24379..24585
 /rpt_family="Li"
 repeat_region 32555..32857
 /rpt_family="Li"
 repeat_region 33985..34281
 /rpt_family="Alu"
 repeat_region 35580..35863
 /rpt_family="Alu"
 complement(31117..31610)
 /rpt_family="Alu"
 repeat_region 38616..38712
 /rpt_family="MIR"
 repeat_region 41075..41331
 /rpt_family="Alu"
 complement(42064..42184)
 /rpt_family="Alu"
 repeat_region 42305..42413
 /rpt_family="MIR"
 complement(42689..42777)
 /rpt_family="Alu"
 repeat_region 43887..44156
 /rpt_family="Alu"
 repeat_region 46963..47240
 /rpt_family="Alu"
 repeat_region 58296..58529
 /rpt_family="Alu"
 complement(51874..52142)
 /rpt_family="Alu"
 repeat_region 52757..52801
 /rpt_family="MERS2"
 repeat_region 58600..58721
 /rpt_family="Li"
 repeat_region complement(5956..59701)

```

repeat_region          repeat_region          105260..105545
/rpt_family="Alu"      /repeat_region          105260..105545
repeat_region          /rpt_family="Alu"      /rpt_family="Alu"
repeat_region          /rpt_family="Alu"      60335
repeat_region          /rpt_family="Alu"      60348..61242
repeat_region          /rpt_family="Alu"      complement(62314..63136)
repeat_region          /rpt_family="Alu"      complement(65228..68344)
repeat_region          /rpt_family="MER20"    /rpt_family="MER20"
STS                   /note="G06727 human STS WI-78431" 69944..70696
repeat_region          /rpt_region          70776..71041
repeat_region          /rpt_region          75220..75585
repeat_region          /rpt_region          complement(77681..77951)
repeat_region          /rpt_region          complement(78093..78393)
repeat_region          /rpt_region          79977..80042
repeat_region          /rpt_region          complement(81069..81163)
repeat_region          /rpt_region          complement(81069..81163)
repeat_region          /rpt_region          81913..81951 human STS CHLC.GGAA 19E07.P17317 clone
repeat_region          /rpt_region          GGA19E07"
repeat_region          /rpt_region          complement(82076..82380)
repeat_region          /rpt_region          82395..82857
repeat_region          /rpt_region          complement(84782..85059)
repeat_region          /rpt_region          complement(86131..86555)
repeat_region          /rpt_region          complement(86698..86772)
repeat_region          /rpt_region          87328..87517
repeat_region          /rpt_region          complement(88025..88300)
repeat_region          /rpt_region          complement(88382..88537)
repeat_region          /rpt_region          complement(9103..92131)
repeat_region          /rpt_region          complement(91607..89870)
repeat_region          /rpt_region          complement(92183..92468)
repeat_region          /rpt_region          complement(92638..92723)
repeat_region          /rpt_region          complement(93254..93542)
repeat_region          /rpt_region          complement(9350..94195)
repeat_region          /rpt_region          complement(94819..94892)
repeat_region          /rpt_region          97105..97412
repeat_region          /rpt_region          complement(101199..101438)
repeat_region          /rpt_region          101753..102038
repeat_region          /rpt_region          complement(102216..102638)
repeat_region          /rpt_region          complement(103792..104101)
repeat_region          /rpt_region          complement(104120..104465)
repeat_region          /rpt_region          104528..10642
repeat_region          /rpt_region          complement(104120..104465)

repeat_region          repeat_region          105260..105545
repeat_region          /rpt_familiy="Alu"      /repeat_region          105260..105545
repeat_region          ORIGIN
repeat_region          Query Match          68.0%; Score 20.4; DB 9; Length 106539;
repeat_region          Best Local Similarity 95.5%; Prd. No. 1..9e+02;
repeat_region          Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
repeat_region          Qy          9 CTTGTAGGAGCCTAGGCTGC 30
repeat_region          Db          69947 CGTGTAGGAGCCTAGGCTGC 69926

RESULT 23
AC125238/C
LOCUS          196622 bp   DNA  PRI 15-OCT-2002
DEFINITION     Homo sapiens BAC clone RP11-711C24 from 2, complete sequence.
AC125238
AC125238..5  GI:22267884
VERSION        HTG.

KEYWORDS       Homo sapiens (human)
SOURCE         Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
               Mammalia; Butheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE     1 (bases 1 to 196622)
AUTHORS        Sulston, J. E. and Waterston, R.
TITLE         Toward a complete human genome sequence
JOURNAL       Genome Res. 8 (11), 1097-1108 (1998)
MEDILINE      99063792
PUBMED        9847074
REFERENCE     2 (bases 1 to 196622)
AUTHORS        Nguyen, C., Bielecki, L. and Schatz, Kamer, K.
TITLE         The sequence of Homo sapiens BAC clone RP11-711C24
JOURNAL       Unpublished (2001)
REFERENCE     3 (bases 1 to 196622)
AUTHORS        Waterston, R. H.
TITLE         Direct Submission
JOURNAL       Submitted (20-JUN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE     4 (bases 1 to 196622)
AUTHORS        Waterston, R. H.
TITLE         Direct Submission
JOURNAL       Submitted (16-AUG-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE     5 (bases 1 to 196622)
AUTHORS        Waterston, R.
TITLE         Direct Submission
JOURNAL       Submitted (15-OCT-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT       On Aug 16, 2002 this sequence version replaced gi:22138709.
COMMENT       Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapis@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0711C24
-----.

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

```

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR:

pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC069314 and AC010138.

A transposon was identified in the finished region of this clone and removed prior to submission. The transposon would insert after base pair 83916 of this sequence.

The sequence from base 156374 to base 156453 was derived from one plasmid subclone.

The sequence from base 156479 to base 156499 was derived from one plasmid subclone.

Polymorphisms have been identified between AC069314, AC010138 and this sequence.

Data from AC069314 and AC010138 was used to finish this clone.

FEATURES	source
	1..196622
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RPCI-711C24"
	/clone.lib="RPCI-11"
	215..236
repeat_region	/rpt_family="(TTTTA)n"
repeat_region	240..520
repeat_region	/rpt_family="Alu"
repeat_region	892..1086
repeat_region	/rpt_family="L1"
repeat_region	1575..1885
repeat_region	/rpt_family="Alu"
repeat_region	1991..2078
repeat_region	/rpt_family="MIR"
repeat_region	2079..2387
repeat_region	/rpt_family="Alu"
repeat_region	2418..2487
repeat_region	/rpt_family="Alu"
repeat_region	2488..2764
repeat_region	/rpt_family="Alu"
repeat_region	2771..2871
repeat_region	/rpt_family="MIR"
repeat_region	3123..3190
repeat_region	/rpt_family="L2"
repeat_region	3296..3599
	/rpt_family="Alu"

```

repeat_region          3612..3919
/rpt_family="Alu"
repeat_region          3926..3988
/rpt_family="MIR"
repeat_region          4192..4493
/rpt_family="Alu"
repeat_region          5165..5556
/rpt_family="ACHobo"
repeat_region          5257..5293
/rpt_family="L1"
repeat_region          5599..5623
/rpt_family="(TG)n"
repeat_region          5624..5748
/rpt_family="Alu"
repeat_region          5752..5929
/rpt_family="L1"
repeat_region          6043..6652
/rpt_family="ACHobo"
repeat_region          6670..6803
/rpt_family="L2"
repeat_region          6804..7097
/rpt_family="Alu"
repeat_region          7098..7192
/rpt_family="L2"
repeat_region          7193..7515
/rpt_family="Alu"
repeat_region          7516..7745
/rpt_family="L2"
repeat_region          7746..8054
/rpt_family="Alu"
repeat_region          8055..8264
/rpt_family="L2"
repeat_region          8265..8565
/rpt_family="Alu"
repeat_region          8566..9242
/rpt_family="L2"
repeat_region          9381..9682
/rpt_family="Alu"
repeat_region          9688..9997
/rpt_family="Alu"
repeat_region          10134..10384
/rpt_family="L2"
repeat_region          10773..11075
/rpt_family="Alu"
repeat_region          11076..11379
/rpt_family="Alu"
repeat_region          11492..11580
/rpt_family="MIR"
repeat_region          11927..12020
/rpt_family="Alu"
repeat_region          12146..12237
/rpt_family="Alu"
repeat_region          12238..12534
/rpt_family="Alu"
repeat_region          12535..12682
/rpt_family="MIR2_Type"
repeat_region          12683..13170
/rpt_family="MIR"
repeat_region          14683..14966
/rpt_family="MIR"
repeat_region          13171..13193
/rpt_family="Alu"
repeat_region          13621..14034
/rpt_family="Alu"
repeat_region          15796..16074
/rpt_family="Alu"
repeat_region          16075..16315
/rpt_family="Alu"

Query Match          68.0$; Score 20.4; DB 9; Length 196622;

```

Best Local Similarity 95.5%; Pred. No. 2e+02; Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 CTTGTCAGGAGCGATGGCTGC 30
Db 127308 CGTGTCAAGCGATGGCTGC 127287

RESULT 24

AC147243 AC147243 181025 bp DNA linear ROD 06-FEB-2004

DEFINITION Mus musculus BAC clone RP23-444P22 from chromosome 6, complete sequence.

ACCESSION AC147243

VERSION AC147243.2 GI:41351711

KEYWORDS MUS

ORGANISM Mus musculus (house mouse)

REFERENCE 1 (bases 1 to 181025)
AUTHORS Shah,N. and Haglund,K.
TITLE The sequence of *Mus musculus* BAC clone RP23-444P22
JOURNAL Unpublished (2001)

REFERENCE 2 (bases 1 to 181025)
AUTHORS Wilson,R.
TITLE Sequencing of *Mus musculus*
JOURNAL Unpublished (2001)

REFERENCE 3 (bases 1 to 181025)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 181025)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (28-JAN-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 181025)
AUTHORS Wilson,R.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-2004) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

COMMENT On Jan 28, 2004 this sequence version replaced gi:38198775.

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@watscan.wustl.edu
----- Summary Statistics
Center project name: M_BA0444P22

SOURCE INFORMATION:
The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa and Minako Tateno in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57Bl/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is overlapped by AC125175.

FEATURES	source	Location/Qualifiers
repeat_region	1..181025	/organism="Mus musculus" /mol_type="genomic DNA" /db_xref="Taxon:10900" /chromosome="6" /map="16" /clone="RP23-444P22" /clone.lib="RPCI-23"
repeat_region	1004..11087	/rpt.family="MIR"
repeat_region	2639..3348	/rpt.family="L1"
repeat_region	3349..4141	/rpt.family="L1"
repeat_region	4744..4894	/rpt.family="Alu"
repeat_region	4924..5110	/rpt.family="B2"
repeat_region	5236..5369	/rpt.family="B2"
repeat_region	5449..5610	/rpt.family="B4"
repeat_region	5756..5897	/rpt.family="Alu"
repeat_region	6146..6318	/rpt.family="B2"
repeat_region	7452..7676	/rpt.family="MER1-type"
repeat_region	9087..9390	/rpt.family="M4LR"
repeat_region	9643..9921	/rpt.family="Alu"
repeat_region	10206..10607	/rpt.family="L1"
repeat_region	10717..10767	/rpt.family="ERV1"
unseq	10723..10775	/note="Sequence derived from one plasmid subclone."
repeat_region	10954..11097	/rpt.family="Alu"
repeat_region	11590..11789	/rpt.family="L1"
repeat_region	13791..14072	/rpt.family="L1"
repeat_region	11900..12056	/rpt.family="MER1-type"
repeat_region	13107..13800	/rpt.family="Alu"
repeat_region	13791..14072	/rpt.family="L1"
repeat_region	16735..16861	/rpt.family="L1"
repeat_region	19455..19597	/rpt.family="RMR19B"
repeat_region	19600..19894	/rpt.family="RMR19B"
repeat_region	21589..22265	/rpt.family="L1"
repeat_region	22270..22768	/rpt.family="L1"
repeat_region	22777..22848	/rpt.family="Alu"

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence, and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

repeat_region 23116..23172
 /rpt_family="ERV1"
 repeat_region 24215..2478
 /rpt_family="ERVK"
 repeat_region 24518..25017
 /rpt_family="ERVK"
 repeat_region 25115..25689
 /rpt_family="L1"
 repeat_region 25738..25880
 /rpt_family="Alu"
 repeat_region 26154..26635
 /rpt_family="L1"
 repeat_region 28759..28874
 /rpt_family="L2"
 repeat_region 29453..29857
 /rpt_family="B4"
 repeat_region 29658..29776
 /rpt_family="L1"
 repeat_region 29776..30334
 /rpt_family="L1"
 repeat_region 30243..31117
 /rpt_family="L1"
 repeat_region 31123..31271
 /rpt_family="B4"
 repeat_region 33546..33538
 /rpt_family="B2"
 repeat_region 34170..34274
 /rpt_family="L1"
 repeat_region 36105..36125
 /rpt_family="L1"
 repeat_region 36505..36680
 /rpt_family="L1"
 repeat_region 37694..38048
 /rpt_family="L1"
 repeat_region 38253..38302
 /rpt_family="L1"
 repeat_region 38940..39031
 /rpt_family="Alu"
 repeat_region 39205..39842
 /rpt_family="L1"
 repeat_region 39765..40442
 /rpt_family="L1"
 repeat_region 40148..40569
 /rpt_family="L1"
 repeat_region 40567..41043
 /rpt_family="L1"
 repeat_region 41076..41529
 /rpt_family="L1"
 repeat_region 42117..42174
 /rpt_family="L1"
 repeat_region 43069..43116
 /rpt_family="ERV1"
 repeat_region 43096..43154
 /rpt_family="ERV1"
 repeat_region 45017..45064
 /rpt_family="L1"
 repeat_region 45065..45912
 /rpt_family="RMR6B"
 repeat_region 45869..46087
 /rpt_family="MBR2_type"
 repeat_region 47846..47944
 /rpt_family="B4"

Query Match 6 GATCTGTGAGCGATGGCTGC 30
 Best Local Similarity 67.3%; Score 20.2; DB 10;
 Matches 88.0%; Pred. No. 2.5e+02; Length 181025;
 22; Conservate 0; Mismatches 3; Indels 0; Gaps 0;

Source 1. 206894 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"

FEATURES

RESULT 25
 AC123799/c
 LOCUS AC123799 Mus musculus chromosome UNK clone RP24-339P7, WORKING DRAFT
 DEFINITION SEQDUNC, 5 unordered pieces.
 AC123799 Mus musculus (house mouse)
 AC123799_2 GI:22138588 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 VERSION 1 (bases 1 to 206894)
 AUTHORS McPherson,J.D. and Waterston,R.H.
 TITLE The sequence of *Mus musculus* clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 206894)
 AUTHORS McPherson,J.D. and Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (01-JUN-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 3 (bases 1 to 206894)
 AUTHORS McPherson,J.D. and Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 COMMENT On Aug 8, 2002 this sequence version replaced GI:21307491.

Genome Center Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site:<http://genome.wustl.edu/gsc/index.shtml>
 Contact: submissions@genome.wustl.edu
 Project Information
 Center project name: M_BB0339P07

Summary Statistics
 Sequencing vector: M13; 0%
 Sequencing vector: plasmid; 100%
 Chemistry: Dye-primer ET; 0% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.99019
 Consensus quality: 205678 bases at least Q40
 Consensus quality: 205903 bases at least Q30
 Consensus quality: 206025 bases at least Q20
 Insert size: 177000; agarose-IP
 Insert size: 206494; sum-of-contigs
 Quality coverage: 19.54 in Q20 bases; agarose-fp
 Quality coverage: 12.13 in Q20 bases; sum-of-contigs

* NOTE: This is a working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 10661: contig of 10661 bp in length
 * 1 10662 10761: gap of unknown length
 * 1 10762 27598: contig of 16337 bp in length
 * 1 27599 27698: gap of unknown length
 * 1 27699 47468: contig of 19770 bp in length
 * 1 47469 47568: gap of unknown length
 * 1 47569 97954: contig of 50386 bp in length
 * 1 97955 98055: gap of unknown length
 * 1 98055 206894: contig of 108840 bp in length.
 Location/Qualifiers


```

repeat_region complement(23503 . 2965)
/rept_family="Trigger6"
repeat_region 3059 . 3493
/rept_family="L1MDa"
repeat_region 3510 . 3851
/rept_family="MTE"
repeat_region 3987 . 4684
/rept_family="L1MDa"
repeat_region 4778 . 4992
/rept_family="L1MDa"
repeat_region 4996 . 5962
/rept_family="L1_MM"
repeat_region 6019 . 6218
/rept_family="L1MDa"
repeat_region 6493 . 6560
/rept_family="AT_rich"
repeat_region complement(6613 . 7554)
/rept_family="MT-int"
repeat_region complement(7579 . 7665)
/rept_family="MTD"
repeat_region 7666 . 7685
/rept_family="(T)n"
repeat_region complement(7686 . 8207)
/rept_family="L1_MM"
repeat_region complement(8208 . 8349)
/rept_family="MTC"
repeat_region complement(8352 . 9904)
/rept_family="L1_RN"
repeat_region complement(9905 . 10147)
/rept_family="MTC"
repeat_region 10217 . 10428
/rept_family="LIMC/D"
repeat_region 11350 . 11390
/rept_family="(TA)n"
repeat_region 12150 . 12211
/rept_family="Lx2"
repeat_region 12712 . 12748
/rept_family="(A)n"
repeat_region 12756 . 12858
/rept_family="Lx2"
repeat_region 12837 . 12983
/rept_family="Lx2"
repeat_region complement(13001 . 13415)
/rept_family="Lx"
repeat_region 13405 . 14051
/rept_family="Lx2"
repeat_region 14062 . 14090
/rept_family="(CAAA)n"
repeat_region 14109 . 14207
/complement(13001 . 13415)
/rept_family="B2_Mm2"
repeat_region complement(14513 . 14629)
/rept_family="B3"
repeat_region 14836 . 14880
/rept_family="(TTTA)n"
repeat_region complement(14882 . 15657)
/rept_family="Lx4"
repeat_region 15700 . 16118
/rept_family="L1"
repeat_region complement(16298 . 17814)
/rept_family="Lx2"
repeat_region complement(17820 . 17915)
/rept_family="Lx2"
repeat_region 17933 . 18942
/rept_family="Lx2"
repeat_region 18955 . 19295
/complement(16298 . 17814)
/rept_family="Lx2"
repeat_region complement(17820 . 17915)
/rept_family="Lx2"
repeat_region 19785 . 19933
/rept_family="MIR"
repeat_region complement(20487 . 21667)
/rept_family="L1_MM"
repeat_region 22750 . 22770
/rept_family="AT_rich"

```

```

repeat_region complement(23503 . 25414)
/rept_family="L1_MM",
repeat_region 25407 . 25714
/rept_family="L1"
repeat_region 25715 . 26058
/rept_family="TAPEY_LTR"

Query Match Score 20 . 2;
Best Local Similarity 88 . 0%; Pred. No. 2 . 5e+02;
Matches 22; Conservate 0; Mismatches 3; Indels 0; Gaps 0;

Qy 6 GATCTGTCAAGGCGATAGGCCTGC 30
Db 89113 GAGCTCTCAGGACATAGGAGC 89089

RESULT 27
AC113935/C
LOCUS AC113935 142422 bp DNA linear PRI 22-MAY-2002
DEFINITION Homo sapiens chromosome 1 clone RP4-706A16, complete sequence.
ACCESSION AC113935
VERSION G1:21070697
HTG.
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 142422)
AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Saenphimachak, C., Phelps, K.A., Buckley, D., Raymond, C. and Haugen, E.D.
TITLE Direct Submission
JOURNAL Unpublished
AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Saenphimachak, C., Phelps, K.A., Buckley, D., Raymond, C. and Haugen, E.D.
COMMENT Center: University of Washington Genome Center
On May 22, 2002 this sequence version replaced gi:19111744.
JOURNAL Box 352145, Seattle, WA 98195, USA
AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Saenphimachak, C., Phelps, K.A., Buckley, D., Raymond, C. and Haugen, E.D.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA
COMMENT Center project name: char-1
Center clone name: RP4-706A16 (SC0534)
SUMMARY Statistics
Sequencing vector: Plasmid; 100% of reads
Chemistry: Dye-terminator ET; 53% of reads
Chemistry: Dye-terminator Big Dye; 47% of reads
Assembly program: Phrap; version 0 . 990319
Consensus quality: 142336 bases at least Q40
Consensus quality: 14421 bases at least Q30
Consensus quality: 142422 bases at least Q20
Insert size: 142422; sum-of-contigs
Quality coverage: 8 . 6x in Q20 bases; sum-of-contigs
OVERLAPPING Sequences:
5 : RP4-564M1 AI035409
3 : RP11-363H12 (OMGC:sc0369) AC093433
SEQUENCE QUALITY Assessment:
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.

```


Query Match	Score 20;	DB 9;	Length 142422;	TITLE JOURNAL	
Best Local Matches	Similarity 82.1%;	Pred. No. 3e+02;	0; Mismatches 5;	Indels 0;	Gaps 0;
Qy	1 TCGAGGATCTGTAGGAGGATAGGCT 28	158498 bp	DNA linear	HTG 13 NOV 2002	COMMENT
Db	111649 TTGTGGATCTGTAGGAGGTGGCT 111622	3 unorderd pieces.	ACCESSION AC098295	GI:24941567 HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.	COMMENT
RESULT 28	AC098295/C	Rattus norvegicus clone CH230-1113,	ORGANISM Rattus norvegicus	Bukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodontia; Sciurognathi; Muridae; Murinae; Rattus.	COMMENT
KEYWORDS	AC098295.4	HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.	VERSION	GI:24941567	COMMENT
SOURCE	AC098295	Rattus norvegicus (Norway rat)	REFERENCE	1 (bases 1 to 158498)	COMMENT
REFERENCE	Authors	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaral-tunge, H.C., Are, J.R., Ayele, M., Banks, T., Barberia, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhey, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Demm, A.L., Ding, Y., Dinh, H.H., Dowthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Barnhart, C., Edgar, D., Edwards, C.C., Elhai, C., Bocatto, M., Palls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Prantz, P., Gabis, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogan, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karatovic, J., Kureshi, A., Khan, U., King, L., Korvar, C., Li, J., Li, Z., Lichatowicz, O., Lieu, C., Liu, J., Liu, W., Loussaged, H., Lopez, R.J., Lu, X., Lucier, A., Lucier, P., Luna, R., Ma, J., Maheeshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Marzke, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newtonson, J., Newtonson, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwoenkwo, S., Oguh, M., Okwounu, G., Oragunye, N., Oriodo, R., Pace, A., Payton, J., Peery, J., Perez, L., Peters, L., Pichens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sison, L., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Tamerisa, A., Tameisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, R., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.P., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.	COMMENT		
REFERENCE	Authors	Worley, K.C.	TITLE JOURNAL	Unpublished	COMMENT
REFERENCE	Authors	Worley, K.C.	TITLE JOURNAL	Direct Submission (23-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	COMMENT
REFERENCE	Authors	Worley, K.C.	TITLE JOURNAL	Submitted (23-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	COMMENT
3 (bases 1 to 158498)	2 CGAGGATCTGTAGGAGGTGGCT 29	3 (bases 1 to 158498)	Query Match	66.7%; Score 20;	COMMENT
3 (bases 1 to 158498)	2 CGAGGATCTGTAGGAGGTGGCT 29	3 (bases 1 to 158498)	Best Local Similarity 82.1%; Matches 23;	Pred. No. 3.1e+02; Mismatches 0;	COMMENT
3 (bases 1 to 158498)	2 CGAGGATCTGTAGGAGGTGGCT 29	3 (bases 1 to 158498)	Indels 0;	Gaps 0;	COMMENT

http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence@genome.wi.mit.edu
Project information
Center project name: L18676
Center clone name: 462_1_5
----- Summary Statistics
Sequencing vector: Plasmid, n/a; 100% of reads
Chemistry: dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0_960731
Consensus quality: 161629 bases at least Q40
Consensus quality: 162397 bases at least Q30
Consensus quality: 162214 bases at least Q20
Insert size: 157000; agarose-fp
Insert size: 162991; sum-of-contigs
Quality coverage: 11.9 in Q20 bases; agarose-fp
Quality coverage: 11.5 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1. 163291; 4936 bp in length
* 1. 4936: contig of 4936 bp in length
* 1. 4936: contig of 100 bp
* 1. 5037: 5036: contig of 622 bp in length
* 1. 5658: gap of 622 bp in length
* 1. 5659: 5758: gap of 100 bp
* 1. 5759: 8865: contig of 3107 bp in length
* 1. 8866: 8965: gap of 100 bp
* 1. 8966: 28315: contig of 1930 bp in length
* 1. 28315: gap of 100 bp
* 1. 28316: 28415: gap of 100 bp
* 1. 28416: 163291: contig of 134876 bp in length.

FEATURES
source
1. 163291
/organism="Mus musculus"
/mol_type="Genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-4621S"
/clone_id="RPCI-24 Male Mouse BAC"
misc_feature
1. 4936
/note="assembly_fragment
clone_end:TP6
vector_side:left"
misc_feature
5037 . 5658
/note="assembly_fragment"
misc_feature
5759 . 8865
/note="assembly_fragment"
misc_feature
8966 . 28315
/note="assembly_fragment"
misc_feature
28416 . 163291
/note="assembly_fragment
clone_end:T7
vector_side:right"
ORIGIN
Birren,B., Nusbbaum,C., Lander,E., Abouel-Hai,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chango,J., Chazarro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Diaz,J.S., Dodges,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galadja,J., Gardyna,S.,
Ginde,S., Gord,S., Goettze,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Keils,C., Lakatos,K.,
Lamazares,R., Landers,T., Lehozky,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McSwain,T., McWatters,K., McPeeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Strange-Thomann,S., Stojanovic,N.,
Strauss,N., Subramanian,A., Talama,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,I., Zimmer,A., and Zody,M.

TITLE JOURNAL
Submitted (3-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE AUTHORS
3. (bases 1 to 163291)
Birren,B., Nusbbaum,C., Lander,E., Abouel-Hai,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gege,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Keils,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Lau,G., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Maneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., O'Connor,T., O'Neill,D., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talama,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,I., Zimmer,A., and Zody,M.

TITLE JOURNAL
Submitted (3-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 7, 2003 this sequence version replaced gi:12381004.
All repeats were identified using RepeatMasker:
Smith,A.F.A., Green,P. (1996-1997)
Db 77991 GAGAACCTCTCAAGGCTATAGCATGC 78018

RESULT 30
AC093433/C AC093433/C linear linear
Query Match 66.7%; Score 20; DB 2; Length 163291;
Best Local Similarity 82.1%; Pred. No. 3.1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 3 GAGGACTTGTCAGGGGATGGCTGC 30
Db 77991 GAGAACCTCTCAAGGCTATAGCATGC 78018

DEFINITION Homo sapiens chromosome 1 clone RP11-363H12, complete sequence.

ACCESSION AC093433 Al390781

VERSION AC093433.2 GI:21954425

KEYWORDS HtG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;

Mammalia; Buthozia; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 172613)

AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Raymond, C., and Haugen, E.D.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 172613)

AUTHORS Haugen, E.D., Olson, M.V., Raymond, C., Clendenning, J., Ivey, R.G. and

TITLE Direct Submission

JOURNAL Submitted (23-AUG-2001) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA

REFERENCE 3 (bases 1 to 172613)

AUTHORS Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Raymond, C., Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M.,

and Haugen, E.D.

TITLE Direct Submission

JOURNAL Submitted (24-JUL-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA

COMMENT On Jul 24, 2002 this sequence version replaced gi:15281284.

Center: University of Washington Genome Center

Center Code: UWGC

Web site: <http://www.genome.washington.edu>

Contact: ungchtg@u.washington.edu

Drafting Center: SC

----- Project Information

Center project name: chr-1

Center clone name: RP11-363H12 (SC0369)

----- Summary Statistics

Sequencing vector: plasmid, L08752; 100% of reads

Chemistry: Dye-terminator Br; 89% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 172300 bases at least Q40

Consensus quality: 172573 bases at least Q30

Consensus quality: 172605 bases at least Q20

Insert size: 172613; sum-of-contigs

Quality coverage: 10.0x in Q20 bases; sum-of-contigs

----- Overlapping Sequences:

5': RP4-706A16 (UWGC:SC0534) AC113935, 49862-bp overlap

3': RP11-375A5 (UWGC:SC0306) AC095030, 1917-bp overlap

----- Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the phrap assembly program. All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest

fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

	BglII	HindIII	EcoRI				
	SqDermP	FngrPrt	SqDermP	FngrPrt	SqDermP	FngrPrt	
	-----	-----	-----	-----	-----	-----	
	2745	2737	8696	8863	10108	10313	
	-----	-----	-----	-----	-----	-----	
	2067	2064	6	<800	6382	6310	
	-----	-----	-----	-----	-----	-----	
	9628	9798	841	847	512	<800	
	-----	-----	-----	-----	-----	-----	
	9759	9798	1070	1010	449	<800	
	-----	-----	-----	-----	-----	-----	
	862	859	14180	13963	3030	3039	
	-----	-----	-----	-----	-----	-----	
	6337	6355	716	<800	5750	5702	
	-----	-----	-----	-----	-----	-----	
	327	<800	234	<800	10429	10313	
	-----	-----	-----	-----	-----	-----	
	4183	4152	3847	3858	2026	2031	
	-----	-----	-----	-----	-----	-----	
	971	994	1973	1947	980	993	
	-----	-----	-----	-----	-----	-----	
	3551	3535	2483	2520	1772	1761	
	-----	-----	-----	-----	-----	-----	
	1007	994	117	<800	2062	2031	
	-----	-----	-----	-----	-----	-----	
	9815	9798	1021	1010	10	<800	
	-----	-----	-----	-----	-----	-----	
	663	<800	4538	4582	9101	8985	
	-----	-----	-----	-----	-----	-----	
	3886	3851	2538	2520	11909	11985	
	-----	-----	-----	-----	-----	-----	
	3878	3851	3451	3433	1585	1550	
	-----	-----	-----	-----	-----	-----	
	1112	1100	11127	11005	1619	1550	
	-----	-----	-----	-----	-----	-----	
	2227	2208	8028	7959	4791	4569	
	-----	-----	-----	-----	-----	-----	
	1600	1547	2736	2786	1197	1170	
	-----	-----	-----	-----	-----	-----	
	3771	3851	3887	3858	2447	2464	
	-----	-----	-----	-----	-----	-----	
	7721	7711	6016	6067	4288	4250	
	-----	-----	-----	-----	-----	-----	
	5015	5005	64	<800	611	<800	
	-----	-----	-----	-----	-----	-----	
	546	<800	1118	1110	640	<800	
	-----	-----	-----	-----	-----	-----	
	13578	13550	2243	2191	11334	11289	
	-----	-----	-----	-----	-----	-----	
	550	<800	213	<800	472	<800	
	-----	-----	-----	-----	-----	-----	
	2797	2842	1486	1465	1309	1311	
	-----	-----	-----	-----	-----	-----	
	3309	3288	953	1010	4116	4050	
	-----	-----	-----	-----	-----	-----	
	1477	1441	2403	2520	2430	2464	
	-----	-----	-----	-----	-----	-----	
	3532	3535	4552	4582	3061	3039	
	-----	-----	-----	-----	-----	-----	
	5347	5354	876	847	5163	5114	
	-----	-----	-----	-----	-----	-----	
	7658	7711	808	847	1532	1550	

FEATURES		Location/Qualifiers	
Source		1..172613	
Query Match	Score 20; DB 9; Length 172613;		
Best Local Similarity	66.7%; Pred. No. 3..1e+02;		
Matches 23; Conservative	0; Mismatches 5; Indels 0; Gaps 0;		
Oy	1 TCGAGGATCTTGTCAGGACGATGGCT 28		
Db	19078 TTGGGATCTGTGAGGATGGCT 19051		
FEATURES			
Source			
RESULT 31			
AC073027_c			
LOCUS	AC073027 173995 bp DNA linear HTG 25-JUN-2000		
DEFINITION	Homo sapiens chromosome 1 clone RP11-295521 map 1, WORKING DRAFT		
SEQUENCE	25 unordered pieces.		
ACCESSION	AC073027		
VERSION	AC073027.2 GI:8705069		
KEYWORDS	HTG; HTGS PHASE1; HTGS_DRAFT.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
1 (bases 1 to 173905)				
Birren,B., Lander,E., Nusbaum,C. and Lander,E.				

REFERENCE	2 (bases 1 to 173995)			
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barnia,N., Basrai,V., Beda,F., Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearall,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreighi,P., FitzHugh,W., Gage,D., Gallegan,J., Gardyna,S., Ginde,S., Goette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kahn,L., Karatas,A., Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehozky,J., Levine,R., Lieu,C., Liu,G., Locke,K., McDonald,P., Marquis,N., McCarthy,M., McBwan,P., McGirk,A., McKernan,K., McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Milenga,Y., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Pierri,N., Piroozi,T., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogoy,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,J., Talamas,J.J., TesFaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Vieil,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainooh,J., Zimmer,A. and ZooY,M.			
JOURNAL	Direct Submission			
-----	Submitted (08-JUN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
COMMENT	On Jun 25, 2000 this sequence version replaced g1:8347970. All repeats were identified using RepeatMasker:			
-----	Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html			
Center: Whitehead Institute/ MIT Center for Genome Research	Genome Center			
Center code: WIBR	Web site: http://www-seq.wi.mit.edu			
-----	Contact: sequence_submission@genome.wi.mit.edu			
-----	Project Information			
-----	Center project name: L10424			
-----	Center clone name: 295_E_21			
-----	Summary Statistics			
Sequencing vector: M13; M77815; 100% of reads				
Chemistry: Dye-terminator Big Dye; 100% of reads				
Assembly program: Phrap; version 0.960731				
Consensus quality: 159136 bases at least Q40				
Consensus quality: 167456 bases at least Q30				
Consensus quality: 170233 bases at least Q20				
Insert size: 171595; sum-of-contigs				
Quality coverage: 3.7 in Q20 bases; sum-of-contigs				
-----	'working draft' sequence. It currently consists of 25 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.			*
-----	This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.			*
-----	1530: contig of 1530 bp in length			*
-----	1630: gap of 100 bp			*
-----	1631: contig of 1470 bp in length			*
-----	3100: gap of 100 bp			*
-----	3101: contig of 1496 bp in length			*
-----	3201: contig of 1496 bp in length			*
-----	4696: contig of 100 bp			*
-----	4697: contig of 1690 bp in length			*
-----	4797: contig of 100 bp			*
-----	6487: contig of 2179 bp in length			*
-----	6586: gap of 100 bp			*
-----	8765: contig of 2179 bp in length			*
-----	8865: gap of 100 bp			*
-----	8866: contig of 2203 bp in length			*
-----	11068: gap of 100 bp			*
-----	11168: gap of 100 bp			*

11169	*	13797:	contig of 2629 bp in length	
13798	*	13897:	gap of 100 bp	
13898	*	16088:	contig of 2201 bp in length	
16089	*	16138:	gap of 100 bp	
16139	*	18174:	contig of 1976 bp in length	
18175	*	18274:	gap of 100 bp	
18275	*	20892:	contig of 2528 bp in length	
20893	*	20902:	gap of 100 bp	
20903	*	23111:	contig of 2209 bp in length	
23112	*	23211:	gap of 100 bp	
23212	*	25268:	contig of 2057 bp in length	
25269	*	25388:	gap of 100 bp	
25369	*	27582:	contig of 2214 bp in length	
27583	*	27612:	gap of 100 bp	
27683	*	31473:	contig of 3791 bp in length	
31474	*	31573:	gap of 100 bp	
31574	*	36258:	contig of 4685 bp in length	
36259	*	36338:	gap of 100 bp	
36359	*	41057:	contig of 4699 bp in length	
41058	*	41157:	gap of 100 bp	
41158	*	49800:	contig of 8643 bp in length	
49801	*	49900:	gap of 100 bp	
49901	*	57393:	contig of 7493 bp in length	
57394	*	5793:	gap of 100 bp	
5794	*	66985:	contig of 9492 bp in length	
66986	*	67085:	gap of 100 bp	
67086	*	82213:	contig of 15128 bp in length	
82214	*	82313:	gap of 100 bp	
82314	*	95892:	contig of 13513 bp in length	
95893	*	95946:	gap of 100 bp	
95947	*	110448:	contig of 14522 bp in length	
110449	*	110548:	gap of 100 bp	
110549	*	129504:	contig of 18954 bp in length	
129603	*	129603:	gap of 100 bp	
129603	*	151126:	contig of 21524 bp in length	
151127	*	151226:	gap of 100 bp	
151227	*	17395:	contig of 22769 bp in length	Location/Qualifiers

FEATURES

```

RESULTS 32
AC026874/c AC026874 203812 bp DNA linear HTG 01-SEP-2000
LOCUS Homo sapiens chromosome 1 clone RP1-627E4, WORKING DRAFT SEQUENCE,
DEFINITION 32 unorderred pieces.
VERSION AC026874
KEYWORDS HTG; HTGS-PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 203812)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 203812)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT On Sep 1, 2000 this sequence version replaced gi:7631093.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH067E04
----- Summary Statistics -----
Sequencing vector: M13, 100%
Sequencing vector: Plasmid, 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-Terminator Big Dye; 0% of reads
Assembly program: Phrap, version 0.930319
Consensus quality: 1863139 bases at least Q40

```

	FEATURES	SOURCE
*	198668 200798: contig of 2131 bp in length	/organism="Homo sapiens"
*	200799 200898: gap of unknown length	/mol_type="genomic DNA"
*	200899 202307: contig of 1409 bp in length	/db_xref="taxon:9606"
*	202308 202407: gap of unknown length	/chromosome="1"
*	202408 203812: contig of 1405 bp in length.	Location/Qualifiers
1.	1..203812	1..
1.	2006: contig of 2006 bp in length	source
2007 2106: gap of unknown length	misc_feature	
2107 4788: contig of 2682 bp in length	misc_feature	
4789 4888: gap of unknown length	misc_feature	
4889 7839: contig of 2951 bp in length	misc_feature	
7840 7939: gap of unknown length	misc_feature	
7940 116227: contig of 3688 bp in length	misc_feature	
116228 11727: gap of unknown length	misc_feature	
11728 16118: contig of 4391 bp in length	misc_feature	
16119 16219: gap of unknown length	misc_feature	
16219 16679: contig of 3461 bp in length	misc_feature	
16679 19779: gap of unknown length	misc_feature	
19779 23680: contig of 3901 bp in length	misc_feature	
23680 23180: gap of unknown length	misc_feature	
23180 282919: contig of 5139 bp in length	misc_feature	
282920 29020: gap of unknown length	misc_feature	
29020 33881: contig of 4861 bp in length	misc_feature	
33881 33980: gap of unknown length	misc_feature	
33981 39616: contig of 5636 bp in length	misc_feature	
39617 39716: gap of unknown length	misc_feature	
39717 46646: contig of 693 bp in length	misc_feature	
46647 46746: gap of unknown length	misc_feature	
46747 52666: contig of 5820 bp in length	misc_feature	
52666 52667: gap of unknown length	misc_feature	
52667 59318: contig of 6652 bp in length	misc_feature	
59319 59418: gap of unknown length	misc_feature	
59419 65916: contig of 6498 bp in length	misc_feature	
65916 66016: gap of unknown length	misc_feature	
66017 72318: contig of 6302 bp in length	misc_feature	
72319 72418: gap of unknown length	misc_feature	
72419 80778: contig of 8360 bp in length	misc_feature	
80779 80878: gap of unknown length	misc_feature	
80879 86913: contig of 6035 bp in length	misc_feature	
86914 87013: gap of unknown length	misc_feature	
87014 94398: contig of 7385 bp in length	misc_feature	
94399 94498: gap of unknown length	misc_feature	
94499 104139: contig of 9641 bp in length	misc_feature	
104140 104239: gap of unknown length	misc_feature	
104240 113608: contig of 9369 bp in length	misc_feature	
113609 113708: gap of unknown length	misc_feature	
113709 114866: contig of 1158 bp in length	misc_feature	
114867 114966: gap of unknown length	misc_feature	
114965 119644: contig of 13056 bp in length	misc_feature	
119645 128023: gap of unknown length	misc_feature	
128023 128122: gap of unknown length	misc_feature	
128123 139097: contig of 10975 bp in length	misc_feature	
139098 139197: gap of unknown length	misc_feature	
139198 149544: contig of 10347 bp in length	misc_feature	
149545 149644: gap of unknown length	misc_feature	
149645 162226: contig of 12582 bp in length	misc_feature	
162227 162326: gap of unknown length	misc_feature	
162327 162337: contig of 15732 bp in length	misc_feature	
162337 178059: gap of unknown length	misc_feature	
178059 181519: contig of 17373 bp in length	misc_feature	
181519 195531: gap of unknown length	misc_feature	
195532 195631: gap of unknown length	misc_feature	
195632 196744: contig of 1113 bp in length	misc_feature	
196745 196844: gap of unknown length	misc_feature	
196845 199867: contig of 1723 bp in length	misc_feature	
199867 199868: gap of unknown length	misc_feature	

misc_feature	RESULT 33	Query Match	Score 20;	DB 2;	Length 203812;	HTG 20-NOV-2002
Best Local Similarity	65.7%					
Matches	82.1%					
Conservative	0;	Pred. No.	3.2e+02;			
REFINEMENT	AC115240	OCTOC	232735 bp	DNA linear		
DEFINITION	AC115240	Rattus norvegicus clone CH230-153F4,		WORKING DRAFT SEQUENCE, 2		
ACCESSION	AC115240	unordered pieces.				
VERSION	AC115240.6	GI:25139449				
KEYWORDS	HTGS PHASEI; HTGS_DRAFT; HTGS_FULLTOP.					
SOURCE	Rattus norvegicus (Norway rat)					
ORGANISM	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
REFERENCE	1 (bases 1 to 232735)	Muzny, D. Marie., Metzger, M. Lee., Abramzon, S., Adams, C., Alder, J., Allen, C., Alien, H., Alsbrooks, S., Amin, A., Anguiano, J., Anyalebechi, V., Aoyagi, A., Ayodele, I. M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, O., Blair, J., Blankenburg, K., Blayth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, Z., Chiu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davis, M.L., Davis, C., Davy Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Desarno, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escott, M., Eugene, C., Evans, C.A., Failes, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Flory, S., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Gady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, K., Harvey, Y., Havlik, P., Hayes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howell, S., Hulyk, S., Hume, J., Idelbier, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowalski, C., Kraft, C.L., Lebow, H., Leyvan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshew, L., Louisaged, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindru, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapia, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Munja, E., Montenayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokelemejeh, O., Okwionu, G., Olarunpunsagon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, B., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shartsen, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smalls, J., Sneed, A., Sodergren, B., Song, X.-Z., Soelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, T., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Wei, X., White, F., Williams, C., Wilson, P., Wierczenko, P., wooden, V.				

misc_feature end sequence:BH336181" 22890. .230652 /note="wgs contig" /note="wgs contig"

misc_feature 230703. 232735 /note="wgs contig"

ORIGIN

Query Match 66.7%; Score 20; DB 2; Length 232735;

Best Local Similarity 82.1%; Pred. No. 3.2e-02; N mismatches 5; Indels 0; Gaps 0;

Matches 23; Conservative 0; N mismatches 5; Indels 0; Gaps 0;

Qy 2 CGAGGATCTTCAGGGATAGGTG 29

Db 218720 CTAGGATTCTCAAGGGATAGGTG 218747

RESULT 34

ACI20701 237445 bp DNA linear HTG 21-SEP-2002

DEFINITION Rattus norvegicus clone CH230-65H6, *** SEQUENCING IN PROGRESS *** , 4 unordered pieces.

ACCESSION ACI20701.4

VERSION GI:23265381

KEYWORDS HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

SOURCE

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 237445)

AUTHORS Muñiz,D.,Marie.,Metzker,M.,Lee.,Abramzon,S.,Adams,C.,Alder,J.,Allen,C.,Allen,H.,Alsbrooks,S.,Amin,A.,Anguiano,D.,Anyalebechi,V.,Ayogai,A.,Ayodeji,M.,Baca,E.,Baden,H.,Baldwin,D.,Bandaranaike,D.,Barber,M.,Barnstead,M.,Benahmed,F.,Biswal,K.,Blair,J.,Blankenburg,K.,Blyth,P.,Brown,M.,Bryant,N.,Buhay,C.,Burch,P.,Burrell,K.,Calderon,B.,Cardenas,V.,Carrier,K.,Cavaroz,I.,Ceasar,H.,Center,A.,Chacko,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,Cleveland,C.,Cockrell,R.,Cox,C.,Coyle,M.,Cree,A.,D'Souza,L.,Davis,M.L.,Davis,C.,Davy-Carroll,J.,de Anda,C.,Dederich,D.,Delgado,O.,Denson,S.,Deramo,C.,Ding,Y.,Dinh,H.,Dirya,K.,Draper,H.,Dugan-Rocha,S.,Dunn,A.,Durbin,K.,Duval,B.,Eaves,K.,Egan,A.,Ebcotto,M.,Eugene,C.,Evans,C.A.,Falls,T.,Fan,G.,Fernandez,S.,Finley,A.,Flagg,N.,Forbes,L.,Foster,M.,Fraser,C.M.,Gabisia,A.,Ganta,R.,Garcia,A.,Garner,T.,Gazza,M.,Georgiadis,E.,Geer,K.,Gill,R.,Grady,M.,Guerra,W.,Guevara,W.,Gumaratne,P.,Haaland,W.,Hamil,C.,Hamilton,C.,Hamilton,K.,Harvey,Y.,Havlik,P.,Hawes,A.,Hernandez,J.,Hernandez,R.,Hines,S.,Hildun,S.L.,Hoogendoorn,A.,Hogues,M.,Hollins,B.,Howells,S.,Hulyk,S.,Hume,J.,Idlebird,D.,Jackson,A.,Jackson,L.,Jacob,L.,Jiang,H.,Johnson,B.,Johnson,R.,Jolivet,A.,Karpathy,S.,Kelly,S.,Kelly,L.,Kovar,C.,Kuan,Z.,King,L.,Kwiat,J.,Liu,W.,Liu,Y.,London,P.,Longmore,S.,Lopez,J.,Liu,J.,Lorensehwa,L.,Loulaged,H.,Loudado,R.J.,Liu,X.,Liu,J.,Mangum,B.,Mapua,P.,Martin,K.,McNeil,T.Z.,Meenin,E.,Nawabiney,S.,McLeod,M.P.,McNeil,T.Z.,Meenin,E.,Milosavljevic,A.,Miner,G.,Minja,E.,Montemayor,J.,Moore,S.,Morgan,M.,Morris,K.,Morris,S.,Munidasan,M.,Murphy,M.,Nair,L.,Nankervis,C.C.,Neal,D.,Newton,N.,Nguyen,N.,Norris,S.,Nwaokelemech,O.,Okwuodu,G.,Olarnpunsagoon,A.,Pal,S.,Parks,K.,Basternak,S.,Paul,H.,Perez,A.,Perez,L.,Martinez,E.,Blomper,F.,Poindexter,A.,Popovic,D.,Primus,F.,Pu,L.-L.,Puazo,M.,Quiroz,J.,Rachlin,E.,Reeves,K.,Regier,M.A.,Reigh,R.,Reilly,B.,Reilly,M.,Ren,Y.,Reuter,S.,Richards,S.,Riggs,F.,Rives,C.,Rodkey,T.,Rojas,A.,Rose,M.,Rose,R.,Ruiz,S.J.,Sanders,W.,Savery,G.,Scherez,S.,Scott,G.,Shatman,S.,Shen,H.,Shetty,J.,Shvartsbeyn,A.,Sison,I.,Sitter,C.D.,Smais,D.,Sneed,A.,Sodergren,R.,Song,X.-Z.,Sorelle,R.,Sosa,J.,Steimle,M.,Strong,R.,Sutton,A.,Svatek,A.,Tabor,P.,Taylor,C.,Taylor,T.,Thomas,N.,Thomas,S.,Tingey,A.,Trejos,Z.,Usmani,X.,Valas,R.,Vera,V.,Villasana,D.,Waldron,L.,Walker,B.,Wang,L.,

REFERENCE 2 (bases 1 to 237445)

AUTHORS Worley,K.C.

JOURNAL Direct Submission

COMMENT Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 237445)

AUTHORS Rat Genome Sequencing Consortium.

JOURNAL Direct Submission

COMMENT Submitted (21-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Sep 21, 2002 this sequence version replaced GI:21908396.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atлас (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

AUTHORS Center: Baylor College of Medicine

JOURNAL Center code: BCM

COMMENT Web site: <http://www.hgsc.bcm.tmc.edu/>

CONTACT: hgsc-hsl@bcm.tmc.edu

Project Information

Center project name: GXOY

Center clone name: CH230-65H6

Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 20781 bases at least Q40

Consensus quality: 21303 bases at least Q30

Consensus quality: 214997 bases at least Q20

Estimated insert size: 233017; sum-of-contigs estimation

Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank/draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 233866: contig of 233866 bp in length

233966: gap of unknown length

234011: contig of 1045 bp in length

235112 236137: contig of 1026 bp in length

236138 236237: gap of unknown length

236238 237445: contig of 1208 bp in length.

FEATURES Location/Qualifiers

source 1. .237445

/note="wgs end extension

Clone_end_77" /note="clone_boundary

clone_end_IT7" /note="clone_boundary

clone_end_IT7" /note="clone_boundary

site:ECORI
end_sequence:BH350813"
complement(23253..233569)
/note="clone boundary
clone end:Sp6
site:ECORI
end_sequence:BH350815"

ORIGIN

Query Match	66.7%	Score 20;	DB 2;	Length 237445;
Best Local Matches	82.1%;	Pred. No. 3.2e+02;	0;	Gaps 0;
Matches	23;	Conservative	0;	Mismatches 5;
Qy	3	GAGGATCTGTCAAGGGCATGGCTGC 30		
Db	23224	GAGGATCTGTCAAGGGCATGGCTGC 23251		

RESULT 35
AC129059 LOCUS Rattus norvegicus clone CH230-1-24M19, WORKING DRAFT SEQUENCE, 2 unorderd pieces.

DEFINITION AC129059

ACCESSION AC129059.4

VERSION GT; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS Rattus_norvegicus (Norway rat)

SOURCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butneria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

ORGANISM 1 (bases 1 to 243439)

REFERENCE MUZNY,D.Marie., METZKER,M.Lee., ABRAMZON,S., ADAMS,C., ALDER,J., ALLEN,C., ALLEN,H., ALISBROOKS,S., AMIN,A., ANGULIANO,D., ANYALEBECHI,V., Aoyagi,A., AYODEJI,M., BACA,E., BADEN,H., BALDWIN,D., BANDARANAIKE,D., BARBER,M., BARNSTEAD,M., BENAHMED,F., BISWALO,K., BLAIR,J., BLURKENBURG,K., BLYTH,P., BROWN,M., BRYANT,N., BUHAY,C., BURCELL,K., CALDERON,B., CARDENAS,V., CARTER,K., CAVAZOS,I., CEASAR,H., CENTER,A., CHACKO,J., CHAVEZ,D., CHAN,G., CHEN,R., CHEN,Y., CHU,J., COLELAND,C., COCKRELL,R., COX,C., COYLE,M., CREE,A., D'SOUZA,L., D'VILA,M.L., DAVIS,C., DAVIS-CARROLL,L., DE ANDA,C., DEDERICHS,D., DELGADO,O., DENSON,S., DERAMO,C., DING,Y., DINH,H., DIVYA,K., DRAPER,H., DUGAN-ROCHA,S., DUNN,A., DURBIN,K., DUVAL,B., EAVES,K., EGAN,A., ESGOTTO,M., EUGENE,C., EVANS,C.A.A., FALLS,T., FAN,G., FERNANDEZ,S., FINLEY,M., FLAGG,N., FORBES,L., FOSTER,M., FRASER,C.M., GABISI,A., GANTA,R., GARCIA,A., GARNER,T., GARZA,M., GBREGEORGIS,E., GEHR,K., GILL,R., GRADY,M., GUERRA,W., GUERRARTE,P., HAALAND,W., HAMIL,C., HAMILTON,C., HAMILTON,K., HARVEY,Y., HAVLAK,P., HAWES,A., HENDERSON,N., HERNANDEZ,J., HERNANDEZ,R., HINES,S., HLADUN,S.J., HODGSON,A., HOGNES,M., HOLLINS,B., HOWELLS,S., HULYK,S., HUME,J., IDLEBIRD,D., JACKSON,A., JACKSON,L., JACOB,L., JIANG,H., JOHNSON,B., JOHNSON,R., JOLIVET,A., KARPATHY,S., KELLY,S., KELLY,S., KHAN,Z., KING,L., KOVAR,C., LIU,J., KRAFT,C.L., LEBOW,H., LEVAN,J., LEWIS,L., LI,Z., LIU,J., LIU,J., LIU,W., LIU,Y., LONDON,P., LONGCRE, S., LOPEZ,J., LORENSENHEWA,L., LOULEGED,H., LOZADO,R.J., LU,X., MA,J., MAHESHWARI,M., MAHMUDARIA,M., MAJILY,K., MANGUM,A., MANGUM,B., MAPUA,P., MARCIN,K., MARTIN,R., MARTINEZ,E., MAWNINY,S., MCLEOD,M.P., MCNEILL,T.Z., MEENEN,E., MELOSAVLJEVIC,A., MINER,G., MINJA,E., MONTEMAYOR,J., MOORE,S., MORGAN,M., MORRIS,K., MORRIS,S., MUNIDASA,M., MURPHY,M., NAIR,L., NANKERVIS,C., NEAL,D., NEWTON,N., NGUYEN,N., NORRIS,S., OLAJUWON,G., OLARMPANSAGOON,A., PAL,S., PARKS,K., PASTERNAK,S., PAUL,H., PEREZ,A., PEREZ,L., PFANNKOCH,C., PLOPPER,P., POINDEXTER,A., POPOVIC,D., PRIMUS,E., PU,L.-L., REILLY,B., QUIROZ,J., RACHLIN,E., REEVES,K., REGIER,M.A., REIGH,R., RIVES,C., RODKEY,T., ROJAS,A., ROSE,R., ROSE,R., RUIZ,S.J., SANDERS,W., SAVERY,G., SCHERER,S., SCOTT,G., SHATTMAN,S., SHEN,H., SHETTY,J., SHVARTSBEN,Y., SITTER,C.D., SMAIS,D., SNEED,A., SODERGREN,B., SONG,X.-Z., SORRIE,R., SOSA,J., STEIMLE,M., STRONG,R., SUTTON,A., SVATEK,A., TABOR,P., TAYLOR,C.,

TITLE
JOURNAL REFERENCE 2 (bases 1 to 243439)
AUTHORS Worley,K.C.
TITLE JOURNAL
JOURNAL OF Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 243439)
REFERENCE Rat Genome Sequencing Consortium.
TITLE JOURNAL
JOURNAL OF Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23264545.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly ('contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

COMMENT
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly ('contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

FEATURES
Source :
1. 242268 contig of 242268 bp in length
* 242269 243439: contig of 1071 bp in length.
Location/Qualifier :
1. 242269 243439: contig of 1071 bp in length.

ORGANISM "Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="Ucaxon:10116"
/clone="CH230-124M19"
347 . 1113
/note="clone_boundary

```

clone_end:T7
site:ECORI
end_sequence:BH329260"
complement(24071..241219)
/note="clone boundary
clone_end:SP6
site:ECORI
end_sequence:BH329263"

RIGIN
Query Match 66.7%; Score 20; DB 2; Length 243439;
Best Local Similarity 82.1%; Pred. No. 3.2e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Y 3 GAGGATCTTGTCAAGGAGCTAGGGCCTGC 30
b 237133 GAGGATCTTGTCAAGGAGCTAGGGCCTGC 237160
ESTRESULT 36
OCUS AC127373 280288 bp DNA linear HTG 25-AUG-2002
DEFINITION Mus musculus chromosome UNK clone RP23-42M8, WORKING DRAFT
SEQUENCE, 10 unordered pieces.
AC127373 GI:21747787
AC127373.1 GI:21747787
HTGS_FULLTOP
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 280288)
McPherson, J.D. and Waterston, R.H.
The sequence of Mus musculus clone
Unpublished
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (14-JUL-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 280288)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web Site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BA0042M08
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-Terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.90319
Consensus quality: 278868 bases at least Q40
Consensus quality: 277312 bases at least Q30
Consensus quality: 277488 bases at least Q20
NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 4562; Contig of 4562 bp in length

FEATURES source
* 4563 4662; Gap of unknown length
* 4663 37815; Contig of 33153 bp in length
* 37816 66758; Contig of unknown length
* 37916 66759 66858; gap of unknown length
* 66859 99532; gap of unknown length
* 99533 141409; Contig of 41777 bp in length
* 141410 141509; gap of unknown length
* 141510 200444; Contig of 58935 bp in length
* 200445 200544; gap of unknown length
* 200545 270671; contig of 70127 bp in length
* 270672 270771; gap of unknown length
* 270772 274472; contig of 3701 bp in length
* 274473 274572; gap of unknown length
* 274573 276805; contig of 2233 bp in length
* 276806 276905; gap of unknown length
* 276906 280288; contig of 3383 bp in length.

Location/Qualifiers
1. .280288
1. .280288
/organism="Mus musculus"
/mol type="Genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/cclone="R223-42M8"

misc_feature
1. .4562
/note="assembly_name:Contig10"
/clone="R223-42M8"
1. .4663
/note="assembly_name:Contig11"
4663-.37815
/note="assembly_name:Contig11"
misc_feature
1. .4663
/note="assembly_name:Contig12"
4663-.37815
/note="assembly_name:Contig12"
1. .47916
/note="assembly_name:Contig13"
47916-.66758
/note="assembly_name:Contig13"
1. .486859
/note="assembly_name:Contig14"
486859-.99532
/note="assembly_name:Contig14"
1. .499633
/note="assembly_name:Contig15"
499633-.141409
/note="assembly_name:Contig15"
1. .500444
/note="assembly_name:Contig15"
500444-.270671
/note="assembly_name:Contig16"
270672-.274472
/note="assembly_name:Contig16"
1. .507072
/note="assembly_name:Contig17"
507072-.276805
/note="assembly_name:Contig17"
1. .5076906
/note="assembly_name:Contig17"
5076906-.280288
/note="assembly_name:Contig17"

ORIGIN
Query Match 66.7%; Score 20; DB 2; Length 280288;
Best Local Similarity 82.1%; Pred. No. 3.3e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
----- SOURCE -----
AP004596 Oceanobacillus iheyensis HTRE831 genomic DNA, section 4/13.
AP004596 BA000028 Oceanobacillus iheyensis HTRE831 linear BCT 17-MAY-2003
AP004596.1 GI:22776527
----- ORGANISM -----
Oceanobacillus iheyensis HTRE831
Oceanobacillus iheyensis HTRE831 Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
1 Lu,J., Nogi,Y. and Takami,H.
----- REFERENCE -----
AUTHORS
TITLE
----- JOURNAL -----
MEDLINE
FEMS Microbiol. Lett. 205 (2), 291-297 (2001)
21623015
-----
```

PUBLMED	11750818	/protein_id="BAC12806_1"
AUTHORS	2 Takami, H., Takaki, Y. and Uchiyama, I.	/db_xref="IGI:22776531"
TITLE	Genome sequence of Oceanobacillus iheyensis isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environments	/translation="MALKSEKIKDCITDVHGIVQVGHHTLYEKINEKDITCTGVTAILPHNRTEEGDTCTINLUVGECDNSYLNISLQATPEHAIAKESASNESSBGAAGTCMOCFGYKGSGIGTASRLQEDVITYLGLGVVSNYNGRKQATFADWEHQETPGCIMMILATDAPUSDRQNLIAKRCAGLGRGTDIGSDAIAFSTANQVSHDSRNHLMSAAYLRDDHPVNOLQFAVEVEITSVLSLQAKTGLRSGHKVEKAPL"
JOURNAL	22220767	complement (1804 . .2790)
MEDLINE	12223376	/gene="OB0851"
PUBLMED	3 (bases 1 to 300050)	complement (2863 . .3459)
AUTHORS	Takami, H., Takaki, Y. and Chee, G.	/gene="OB0851"
TITLE	Direct Submission	/note="CDS_ID OB0851"
JOURNAL	Submitted (26-DEC-2001) Hideyo Takami, Japan Marine Science and Technology Center, Deep-sea Microorganisms Research Group; 2-15 Natsushima-cho, Yokosuka, Kanagawa 237-0061, Japan	/codon_start=1
FEATURES	(E-mail:takami@jamstec.go.jp);/amstec-e/bio/DEEPSTAR/FResearch.html, URL:http://www.jamstec.go.jp/jamstec-e/bio/DEEPSTAR/FResearch.html, Tel:81-468-67-9643, Fax:81-468-67-9645)	/transl_table=11
source	Location/Qualifiers	/product="hypothetical protein"
CDS	1 . .300050	/db_xref="IGI:22776532"
gene	/organism="Oceanobacillus iheyensis HT0831"	/translation="MKKKWVKPITLGTLLFGPLFYDQEVEPKPETNATKNSDENEEKEVYKDPKISEBTKENANPFGDNVSASMNRBLIKYHMFHQKVEAQSKNGFYEITDORIAFQEBSENITDSNAFEVERASERNILDWSNSGDFSQVSDHNTLYRNSEHDBDRGRATGATDDEBNKTFENKOFSKIRGJNRFDE"
CDS	/mol_type="genomic DNA"	/codon_start=1
gene	/strain="HT0831"	/transl_table=11
CDS	/note="CDS_ID OB0847"	/product="cytochrome c biogenesis (thiol-disulfide interchange protein)"
gene	/db_xref="taxon:221109"	/protein_id="BAC12806_1"
CDS	/note="halotolerant and alkaliphilic"	/db_xref="IGI:22776533"
gene	154 . .441	/translation="MKILKTVLIIISMLPIMTVAAGTEVGERAPPDEFLKTIDQQLRLSDPKGERVLINFWTTWCPPCRQMPDMQRFYQDQPNLAVNLTDENNKEQVFRFSQEELETFPLILDERKEVSKAYRISPIPTYMIDSEGTRHKSYGALTVEQMAEVNKME"
CDS	/gene="OB0847"	/codon_start=1
gene	154 . .441	/transl_table=11
CDS	/gene="OB0847"	/product="hypothetical conserved protein"
gene	/note="CDS_ID OB0847"	/protein_id="BAC12804_1"
CDS	/codon_start=1	/db_xref="IGI:22776534"
gene	/transl_table=11	/translation="NDKGCTIRTIVLVALTNOELITAGYHPIPCTQELWGEILSSIFTIVATUQNNVYVTKQQKHHQVHEHQLAK"
CDS	/product="hypothetical conserved protein"	/codon_start=1
gene	/protein_id="BAC12804_1"	/product="N-acetyl muramoyl-L-alanine amidase (sporulation mother cell wall hydrolase)"
CDS	/db_xref="IGI:22776530"	/protein_id="BAC12810_1"
gene	/note="CDS_ID OB0848"	/db_xref="IGI:22776535"
CDS	/codon_start=1	/translation="MKRIFIDGHGNDPGAMGNGLOEKOINNDIALRKEYLLQFTGHEVRUSRETDKTVSLSQRTMANWQADYLVLVHSNAGGTGTFESYTFNGYSGRAETHRLSQLVHQSVNNTGYDGMKEANHMLRESQMPAVLTENGFDYREDAKLKSSFLQSSTARSHATQVLLDGEAGASQGVLELAUDWTINSPEWEDKAVIVHKGEVFTVVKDKDPVQDGMYQIKSLUYITANKVYTRYTSTSCKTSVLS"
gene	/transl_table=11	/gene="OB0854"
CDS	/product="hypothetical conserved protein"	/note="CDS_ID OB0854"
gene	/protein_id="BAC12804_1"	/codon_start=1
CDS	/db_xref="IGI:22776529"	/product="hypothetical conserved protein"
gene	/note="CDS_ID OB0849"	/protein_id="BAC12803_1"
CDS	/codon_start=1	/db_xref="IGI:22776534"
gene	/transl_table=11	/translation="NDKGCTIRTIVLVALTNOELITAGYHPIPCTQELWGEILSSIFTIVATUQNNVYVTKQQKHHQVHEHQLAK"
CDS	/product="hypothetical conserved protein"	/codon_start=1
gene	/protein_id="BAC12805_1"	/product="D-amino peptidase"
CDS	/db_xref="IGI:22776530"	/note="CDS_ID OB0855"
gene	/note="CDS_ID OB0850"	/codon_start=1
CDS	/codon_start=1	/transl_table=11
gene	/transl_table=11	/product="hypothetical protein"

/protein_id="BAC12811_1"
 /db_xref="GI:22776536"
 /translation="MKTLLGVVITALIMNGMASFYDQMNNEINTWIFYLVTSMILLL
 VLVIKEWFLRKSTKDKDINE"
gene
 5593 . '7476
 /gene="OB0856"
 5593 . '7476
 /gene="OB0856"
 /note="CDS ID OB0856"
CDS
 /codon_start=1
 /transl_table=11
 /product="ABC transporter ATP-binding protein"
protein_id="BAC12812_1"
 /db_xref="GI:22776537"
 /translation="MLMIDNISKTYGCDKTLFENQITATIGERKRIGLIGVNCTGKSFL
 KVIAGMDSPDODKDIKSKDYIIEBYTQETNLQEDITVLDOLYYGDALIMVMROYEQAA
 LIDDLENSPEADALQQLQMDPMTNHLHDTPTEAWTAKTLYMQYSGSVLLVTHDRYFLNR
 QKRIIAKALIOPADLIMDPMNTNHLHDTPTEAWTAKTLYMQYSGSVLLVTHDRYFLNR
 VTNRIFELDKENLYTAGNEYELFELEKAERSELERNEOKHONTLKRELAWLKGAKA
 RSTKQRARVERTEBEMDKTFDKEKEVFKQRLGNDVLEIDGYKSGQNQTVLJSN
 PSQLIIPEDRIGLIGNGTGKTMNMAKTEPDQGYVKGETVLGYTQQGEERBLD
 DQORILEYKQVAEVHTKGGEVITAQMFLFRACOMSFIRLSCGGRRLYL
 KVLMKPENVLFLDEPDINDLDTLSVUDLFLPGVIVTSHDFRNLDRVNKLIVF
 KGNGNTDFYGYTPELQKQRAEKLOQTPKQSSTTSQEKKKLSSYNEKRENDITI
 EBEITKLEEEVTEQIAVTDAGSDSERVOLEYAQAVELAEEWMARWEALSIVQE
gene
 7769 . '8821
 /gene="OB0857"
 7769 . '8821
 /gene="OB0857"
 /note="CDS ID OB0857"
CDS
 /codon_start=1
 /transl_table=11
 /product="hypothetical conserved protein"
protein_id="BAC113_1"
 /db_xref="GI:22776538"
 /translation="MOBIYSDVITFERSHYOFGKMOGESIKRSFYMENRKQWKLRRP
 RPTINIDEVKSMINTYAPGIVNLELEGKDGTEWSLSDDTQEAQGRINLDSGGCVFT
 SSNNMIRNYDYPKTPKTYEGRLLFQPTDGLTGTGAGGTAGGCT 71034

Query Match 66.7% Score 20; DB 1; Length 300050;
 Best Local Similarity 82.1%; Pred. No. 3.3e+02; Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TCGAGGATCTTGTCAGGACGATAAGCT 28
 Db 71007 TCAATGTTCTTGTAGCACCGATAAGCT 71034

RESULT 38
 BX005159/c
LOCUS BX005159 312728 bp DNA linear HTG 08-JAN-2003
DEFINITION Homo sapiens chromosome 1 Clone RP11-619A13.
VERSION BX005159.2 GI:27650315
KEYWORDS HTG; HTGS PHASE2; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
MATERIALS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE Hall,R. Direct Submission
JOURNAL Submitted (07-JAN-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail requests: humquery@sanger.ac.uk
COMMENT On Jan 11, 2003 this sequence version replaced gi:27368635.
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: bA619A13
 ----- Summary Statistics
 Assembly program: XGAP4 ; version 4.5
 Chemistry: Dye-terminator ; 100% of reads
 Consensus quality: 167037 bases at least Q40
 Consensus quality: 167127 bases at least Q30
 Consensus quality: 167156 bases at least Q20
 Insert size: 312728; sum-of-contigs
 Insert size: 16110; 3.9% error; agarose-fp
 Quality coverage: 4.11x in Q20 bases; sum-of-contigs Quality coverage: 8.01x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter.
 This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
 * the 1 312728: contig of 312728 bp in length.
FEATURES
source /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-619A13"
 /clone.lib="RPCI-11.3"
misc_feature 1 . 312728
 /note="assembly_fragment:01914
 clone.end:SP6
 vector.side:left
 clone.end:RT
 vector.side:right"
ORIGIN
 Query Match 66.7% Score 20; DB 2; Length 312728;
 Best Local Similarity 82.1%; Pred. No. 3.3e+02; Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 TCGAGGATCTTGTCAGGACGATAAGCT 28
 Db 19078 TTCTGGACTCTGTGAGGTGATTGGCT 19051

RESULT 39
 AC136270/c
LOCUS AC136270 331221 bp DNA linear HTG 23-NOV-2002
DEFINITION Rattus norvegicus clone CH230-46911, *** SEQUENCING IN PROGRESS

ACCESSION AC136270 *** 12 unorderd pieces.
VERSION AC136270.2 GI:25188310
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
MATERIALS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Rattus.
REFERENCE 1 (bases 1 to 312721)
AUTHORS Muzny, D. Marie, Metzker, M. Lee, Abramzon, S., Adams, C., Alder, J., Allen, C., Allier, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyabechi, V., Ayodeji, A., Baden, H., Baca, R., Baden, H., Baldwin, D., Bandaraanake, D., Barber, M., Barnstead, M., Benahmed, F., Biswalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Calderon, E., Cardenas, V., Carter, K., Cavares, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Andra, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Durbin, K., Dunn, A., Durbin, B., Eaves, K.,

Egan,A., Escott,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabis,A., Gant,R., Garcia,A., Garner,T., Garza,M., Geboegeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guivara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hines,S., Hladun,S.L., Henderson,N., Hernandez,J., Hernandez,R., Hollins,B., Howells,S., Jackson,A., Hogan,M., Hollins,B., Howells,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Khan,Z., Koivar,C., King,L., Lewis,J., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Louis,L., Lozado,R.J., Lu,X., Ma,J., Maheshwari,L., Malmud,M., Mahmudarre,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,B., Mawhinney,S., McLeod,M.P., McNeil,T.Z., Meienen,B., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidesa,M., Murphy,M., Nair,L., Narkervis,C., Neal,D., Newton,N., Norris,S., Okwonu,G., Olarmpusnagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,I.D., Primus,B., Pu,L.-L., Puaico,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,P., Rives,C., Rodkey,T., Rojas,A., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sned,A., Sodegren,E., Song,X.-Z., Sorela,J., Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vira,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczky,R., Wooden,H., Worley,K., Wright,D., Wright,Y., Wu,J., Yakuob,Y., Yoon,L., Yoon,V., Yu,P., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G., and Gibbs,R.A.

TITLE

Unpublished

(bases 1 to 331221)

Rat Genome Sequencing Consortium

JOURNAL

Direct Submission

Submitted (31-OCT-2002)

Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 331221)

Rat Genome Sequencing Consortium

REFERENCE

Direct Submission

Submitted (23-NOV-2002)

Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 23, 2002 this sequence version replaced gi:24431541.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (an 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.edu/>Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KBUS

Center Clone name: CH230-46911

----- Summary Statistics

Assembly program: phrap; version 0.990329
 Consensus quality: 257314 bases at least Q40
 Consensus quality: 26539 bases at least Q30
 Consensus quality: 271049 bases at least Q20
 Estimated insert size: 267830; sum-of-contigs estimation
 Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see <http://www.hgsc.bcm.tmc.edu/dccs/genbank/draft-data.html>)
 * NOTE: This sequence may represent more than one clone.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

<p>** 1 4872: contig of 4872 bp in length ** 4873 4972: gap of unknown length ** 25330: contig of 30358 bp in length ** 25331 25430: gap of unknown length ** 25431 112631: contig of 87201 bp in length ** 112632 112731: gap of unknown length ** 112732 262442: contig of 149711 bp in length ** 262443 262542: gap of unknown length ** 306832 306833: contig of 44280 bp in length ** 306833 306922: gap of unknown length ** 306923 317634: contig of 10692 bp in length ** 317635 317734: gap of unknown length ** 317735 319314: contig of 1600 bp in length ** 319315 319414: gap of unknown length ** 319415 320445: contig of 1035 bp in length ** 320450 320549: gap of unknown length ** 320550 322012: contig of 1463 bp in length ** 322013 322112: gap of unknown length ** 322113 324044: contig of 1932 bp in length ** 324045 324145: gap of unknown length ** 324145 328550: contig of 4406 bp in length ** 328551 331221: contig of 2571 bp in length.</p>	<p>FEATURES source Location/Qualifiers 1. .331221</p>
--	---

/organism="Rattus norvegicus"
 /mol type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-46911"
 misc_feature
 25431..26656
 /note="wgs contig"
 misc_feature
 77244..78677
 /note="wgs contig"
 misc_feature
 112732..114588
 /note="wgs contig"
 misc_feature
 15245..154397
 /note="wgs contig"
 misc_feature
 158417..160948
 /note="wgs contig"
 misc_feature
 172688..174088
 /note="wgs contig"
 misc_feature
 202512..204106
 /note="wgs contig"
 misc_feature
 244136..245292
 /note="wgs contig"
 misc_feature
 262543..263907
 /note="wgs contig"

ORIGIN

Query Match Score 66.7%; Best Local Similarity 82.1%; Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 2 CGAGGATCTTGTCAAGCGCATGGCTG 29

Db 197084 CTAGGATTCTGTCAGAGGGATAGGCTG 197057

RESULT 40

AY057108/C

LOCUS AY057108 5359 bp DNA linear PLN 24-APR-2002

DEFINITION Hypocrea jecorina QM9414 transcriptional adaptor-like protein (ada2), galactose 1-phosphate uridylyltransferase (gal7), and transmembrane domain-containing protein (erg28) genes, complete cds.

ACCESSION AY057108

VERSION AY057108.2 GI:18101631

SOURCE Hypocrea jecorina

ORGANISM Hypocrea jecorina

REFERENCE 1 (bases 1 to 5359) Hypocreomycetidae; Pezizomycotina; Sordariomycetes; Hypocreales; Hypocreaceae; Hypocreales.

AUTHORS Seibold, B., Hofmann, G. and Kubicek, C.P.

TITLE lacticose metabolism and cellulase production in Hypocrea jecorina: the gal7 gene, encoding galactose-1-phosphate uridylyltransferase, is essential for growth on galactose but not for cellulase induction.

JOURNAL Mol. Genet. Genomics 267 (1), 124-132 (2002)

JOURNAL 219119477

PUBLMED 119119723

REFERENCE 2 (bases 1 to 5359)

AUTHORS Seibold, B., Hofmann, G. and Kubicek, C.P.

TITLE Direct Submission

JOURNAL Submitted (09-JAN-2002) IBTM, TU Vienna, Getreidemarkt 9, Vienna

JOURNAL 1060, Austria

REFERENCE 3 (bases 1 to 5359)

AUTHORS Seibold, B., Hofmann, G. and Kubicek, C.P.

TITLE Submitted (09-JAN-2002) IBTM, TU Vienna, Getreidemarkt 9, Vienna

JOURNAL 1060, Austria

COMMENT Sequence update by submitter

FEATURES On Jan 9, 2002 this sequence version replaced gi:16209627.

source Location/Qualifiers

1. .5359

/organism="Hypocrea jecorina"

/mol_type="Genomic DNA"

/strain="QM9414"

/db_xref="taxon:51453" complement (>361..>2015)

/gene="ada2"

complement (<361..>2015)

/gene="ada2"

/product="transcriptional adaptor-like protein"

/protein_id="AAU14200_2"

/db_xref="GI:18101632" translation="MGVTRKCTATRGEGGVYKVCDVCSIDLTSTVRCADPAGSDF

DCCVPCPKGCKGPSPRNADPATPRVEQNSDIFDRGPGDPLLLEGARIGGSW

AJIADHGFFRKEDYRDHYKTYIBSPNFDPLPERSPHNELANEIPIRFQARKCR

RIEERREASKNPAPALOFPKTKTAVPSCHETQGYMPGRLFETETANRAEAEVOLMOF

DGGDGINKPTCELEPEMELKLUTVMELYNCRITQVERKVKTFENALDFTNTKOBKR

KKEERDILQRKPFARIIMKHDDEDTQGQDLELNROIAAQLOQWSRSKKGIDURSG

EKYEAEKAARIOKAIPMGSDRERASSORSKAOPPEPPSGASLLVAPLPTEAQOA

NCBANGEQDKGKATLNGQNSDINGHSPIAKQKTFQPSGVPLQLTENADPHILLPEE

ARLICEVRLQPPKPYLMIKE1KE1KGNTFLKKOAKE1CRLDSQKGARLFDPFSNA

GWVGKA"

/note="gal7">4505

/join(<3150..>4505)

/gene="gal7"

/mRNA .3250,3325..3376,3444..3579,3646..>4505)

/gene="gal7"

/product="galactose-1-phosphate uridylyltransferase"

/join(3150..3250,3325..3376,3444..3579,3646..4505)

/gene="gal7"

/mRNA

/CDS

/EC number="2.7.7.12"

/note="Gal7"

/codon_start=1

/product="galactose-1-phosphate uridylyltransferase"

/protein_id="AAU14201_1"

/db_xref="GI:16209629"

/translation="MFDKILDDISHRRYNNPLTDMSLIVSPHTKRPWQGOBGAATT

LRAOGYKGVCYLTFRPKHNTYQNDTSAVNDYQATVAFVNDYQATVAFVNDYQATV

LQLPISKEAEPVKONYMQIFENKGAMGCSNPHPCQAWTTSTMPEEFGKELVOM

AKYRQORGHLLADYIKLLELAKEERYVQNDSFVVVCWPAIWPFVVLPKRHVAL

PPLLRGATVKKFVGPBMFAAQDITPEQATIRLRACDGELYRNLKSL"

complement (<4623..>5208)

/complement(join(<4623..>5208))

/gene="erg9-8"

/product="transmembrane domain-containing protein"

/complement(join(4623..5072,5146..5208))

/gene="erg9-8"

/note="erg9-8"

/product="Brg28p similar to the Brg28p of Saccharomyces cerevisiae"

/codon_start=1

/product="transmembrane domain-containing protein"

/protein_id="AAU14202_2"

/db_xref="GI:18101633"

/translation="MDLRLKTLPPARGLIPYMLLVSIGNSLQAYATLHFSSRRVY

NGREFRNPKLUPPASAKFNPKIDSPNKLVHAONDPKATDQTLPLAGLUGFTWLTCIVR

CYAATHLHLGPVYNIAIWTVVALGHFASEBLVFKSMTGLQIPFTLATTALIWMP

LVRSHYVEFB"

ORIGIN

Query Match

Best Local Similarity

Matches 22; Conservative

0; Mismatches 4;

Indels 0; Gaps 0;

Qy 1 TCGAGGATCTTGTCAGGAGCAGTAGG 26

Db 3169 TCGAGGATCTTGTCAGGAGCAGTAGG 3144

Search completed: January 7, 2005, 11:33:35

Job time : 1039.21 secs

gene

/mRNA

/CDS

/note="Ada2"

/codon_start=1

/product="transcriptional adaptor-like protein"

/protein_id="AAU14200_2"

/db_xref="GI:18101632" translation="MGVTRKCTATRGEGGVYKVCDVCSIDLTSTVRCADPAGSDF

DCCVPCPKGCKGPSPRNADPATPRVEQNSDIFDRGPGDPLLLEGARIGGSW

AJIADHGFFRKEDYRDHYKTYIBSPNFDPLPERSPHNELANEIPIRFQARKCR

RIEERREASKNPAPALOFPKTKTAVPSCHETQGYMPGRLFETETANRAEAEVOLMOF

DGGDGINKPTCELEPEMELKLUTVMELYNCRITQVERKVKTFENALDFTNTKOBKR

KKEERDILQRKPFARIIMKHDDEDTQGQDLELNROIAAQLOQWSRSKKGIDURSG

EKYEAEKAARIOKAIPMGSDRERASSORSKAOPPEPPSGASLLVAPLPTEAQOA

NCBANGEQDKGKATLNGQNSDINGHSPIAKQKTFQPSGVPLQLTENADPHILLPEE

ARLICEVRLQPPKPYLMIKE1KE1KGNTFLKKOAKE1CRLDSQKGARLFDPFSNA

GWVGKA"

/note="gal7">4505

/join(<3150..>4505)

/gene="gal7"

/mRNA .3250,3325..3376,3444..3579,3646..>4505)

/gene="gal7"

/product="galactose-1-phosphate uridylyltransferase"

/join(3150..3250,3325..3376,3444..3579,3646..4505)

/gene="gal7"

Copyright (c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6

OM nucleic - nucleic search, using SW model

Run on: January 7, 2005, 10:58:39 ; Search time 930.789 Seconds
(without alignments)

1371.762 Million cell updates/sec

Title: US-09-786-502A-7

Perfect score: 27

Sequence: 1 gggccgaaatggatgtatct 27

Scoring table: IDENTITY_NUC

Gapop 10.0 - Gapext 1.0

Searched: 4526729 seqs, 2364489745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```

1: GenEmbl:*
2: gb_ba:*
3: gb_htg:*
4: gb_om:*
5: gb_ov:*
6: gb_dat:*
7: gb_ph:*
8: gb_dl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	27	100.6	27	6	BD272284	BD272284 Fused receptor specific to prostate-specific membrane antigen and its use.
c	19.2	71.1	189912	2	AC102434	AC102434 Fused receptor specific to prostate-specific membrane antigen and its use.
3	19.2	71.1	214580	10	AC102369	AC102369 Fused receptor specific to prostate-specific membrane antigen and its use.
4	19.2	71.1	245862	2	AC123598	AC123598 Fused receptor specific to prostate-specific membrane antigen and its use.
5	19	70.4	542	9	HUMCD282	HUMCD282 Fused receptor specific to prostate-specific membrane antigen and its use.
6	19	70.4	663	6	AX695883	AX695883 Fused receptor specific to prostate-specific membrane antigen and its use.
7	19	70.4	705	9	AU517504	AU517504 Fused receptor specific to prostate-specific membrane antigen and its use.
8	19	70.4	1514	6	BD267803	BD267803 Methods a
9	19	70.4	1514	6	I07254	I07254 Sequence 11
10	19	70.4	1514	6	AR380771	AR380771 Sequence 11
11	19	70.4	1514	6	BD015187	BD015187 TtLSa_ce1
12	19	70.4	1514	9	HUNSPCTS	HUNSPCTS TtLSa_ce1
13	19	70.4	3803	6	AX557214	AX557214 Sequence
14	19	70.4	3804	6	AX695882	AX695882 Sequence
15	19	70.4	3806	6	CQ715869	CQ715869 Sequence
16	19	70.4	3806	6	AX285559	AX285559 Sequence
17	19	70.4	51365	6	AX695881	AX695881 Sequence
18	19	70.4	66887	9	AP411057	AP411057 Homo sapi
19	19	70.4	88624	2	AC020676	AC020676 Homo sapi

ALIGNMENTS

RESULT 1	BD272284	27 bp linear
LOCUS	DEFINITION	DNA
ACCESSION	VERSION	Fused receptor specific to prostate-specific membrane antigen and its use.
KEYWORDS	SOURCE	BD272284_1 GI:33082052
ORGANISM	SYNTHETIC CONSTRUCT	JP 2002524081-A/7
REFERENCE	ARTIFICIAL SEQUENCES.	1 (bases 1 to 27)
AUTHORS	TITLE	Sadeain,M., Bande,N.H. and Gong,M.
JOURNAL	COMMENT	Fused receptor specific to prostate-specific membrane antigen and its use.
SLOAN KETTERING INSTITUTE FOR CANCER RESEARCH	COMMENT	Patent: JP 2002524081-A/7 06-AUG-2002;
OS	COMMENT	Synthetic construct
PN	PN	JP 2002524081-A/7
PD	PD	06-AUG-2002
PP	PP	JP 20000568998
PR	PR	03-SEP-1999 US 60/09138
PI	PI	MICHEL SADELAIN, NEIL H BANDER, MICHAEL GONG
PC	PC	C12N15/00, A61K35/26, A61K35/76, A61K39/395, A61K48/00, A61P35/00,
PC	PC	C07K19/00
PC	PC	C12N5/10//C07K16/44, C12N15/00, C12N5/00
CC	CC	Fused receptor specific to prostate-specific membrane antigen and its use.
FH	FT	Key source
FT	FT	/organism='Synthetic construct'.
FT	FT	Location/Qualifiers
FT	FT	1..27
FT	FT	/organism='synthetic construct'
FT	FT	/mol_type='genomic DNA'
FT	FT	/db_xref='taxon:32630'
ORIGIN	FT	Origin

Query Match 100.0%; Score 27; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.01%;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGCGGAAATTGAAGTTATGTTATGTCCT 27
 Db 1 GGGCGGAAATTGAAGTTATGTCCT 27

RESULT 2
 AC102434/c LOCUS AC102434 DNA linear HTG 25-MAR-2004
 DEFINITION Mus musculus chromosome 12 clone RP24-120L6 map 12; *** SEQUENCING
 AC102434
 VERSION GI:45736589
 KEYWORD HTGS PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
 SOURCE HTGS (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Muridae; Mus.
 Birren,B., Nusbaum,C. and Lander,E.
 REFERENCE 1 (bases 1 to 189912)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 DeArellano,K., Dewart,K., Diaz,J.S., Dodge,S., Faro,S.,
 Gallegos,R., Haford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kelis,I., Latocque,K.,
 Lehozky,J., Levine,R., Liu,G., Major,J., Marquis,N., Matthews,C.,
 MacLean,C., McDonald,P., McFarlane,P., McKernan,K., McPheeters,R., Meldrim,J.,
 McCarthy,M., McSwain,P., Menger,V., Murphy,T., Naylor,J., Nguyen,C.,
 Neenes,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 O'Halloran,T., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spence,J., Strange-Thomann,N., Stojanovic,N.,
 Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.-J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission
 JOURNAL Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 189912)
 AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelelleil,A., Allen,N.,
 Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
 Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J.,
 Choeipel,Y., Collymore,A., Cook,P., Corum,B.,
 DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
 Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
 Galagan,J., Gardyne,S., Graham,L., Grand-Pierre,N.,
 Hagopian,D., Hagos,B., Hall,J., Horton,L., Huime,W., Iliev,I.,
 Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., Landers,T.,
 Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R.,
 MacLean,C., McDonald,P., Major,J., Manning,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
 Phunkhang,F., Pierre,N., Rachupa,A., Ramasamy,U., Raymond,C.,
 Retta,R., Riese,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Smith,C., Spence,J., Strange-Thomann,N.,
 Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
 Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
 Zimmer,A. and Zody,M.

TITLE Direct Submission
 JOURNAL Submitted (25-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Mar 25, 2004 this sequence version replaced gi:30017683.

All repeats were identified using RepeatMasker:
 Smit, A.P.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www.seq.wi.mit.edu
 Contact: submissions@broad.mit.edu
 ----- Project Information
 Center project name: L18789
 Center clone name: 120_L_6

NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 60150: contig of 60150 bp in length
 * 60151 60250: gap of 100 bp
 ** 60251 94286: contig of 34036 bp in length
 ** 94287 94387: gap of 100 bp
 ** 94387 98686: contig of 4300 bp in length
 ** 98687 98687: gap of 100 bp
 ** 98787 142004: contig of 43218 bp in length
 * 142005 142104: gap of 100 bp
 * 142105 189912: contig of 47808 bp in length
 1. 189912 Location/Qualifiers
 source
 /organism="Mus musculus"
 /mol_type="Genomic DNA"
 /db_xref="Itaxon:10090"
 /chromosome="12"
 /map="1.2"
 /clone="RP24-120L6"
 /clone_id="RPCI-24 Male Mouse BAC"
 ORIGIN
 Query Match Score 71.1%; Best Local Similarity 87.5%; Matches 21; Conservation 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 3 GCGCCGATTGAGTGTATGATGATC 26
 Db 106805 GGCTCTAATTGGAGATATGATGATC 106782
 RESULT 3 AC102369
 LOCUS AC102369 DNA linear ROD 03-JUN-2004
 DEFINITION Mus musculus chromosome 12, clone RP23-246F14, complete sequence.
 ACCESSION AC102369
 VERSION AC102369-21 GI:48059446
 HTG.
 KEYWORDS SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCES 1 (bases 1 to 21580)
 AUTHORS Birren,B., Colangelo,M., Collins,S., Collymore,A., Chazaro,B.,
 Choeipel,Y., Dehrellano,K., Devor,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Grand-Pierre,N.,
 Ginde,S., Gord,S., Goette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,

Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPhee, R., Meldrum, J.,
 Maneus, L., Mihova, T., Mlenga, V., Murphy, T., Navljar, J., Nguyen, C.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunckhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rogov, P.,
 Ronan, J., Roseti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
 Seaman, S., Severny, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
 Straus, N., Subramanian, A., Talmas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Triaglio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zemberk, L., Zimmer, A., and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 24580)
AUTHORS Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V.,
 Bloom, T., Boguslavskiy, L., Boukgalter, B., Camara, J., Chang, J.,
 Choepel, Y., Collimore, A., Cook, A., Cooke, P., Corum, B.,
 DeAreliano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L.,
 Erickson, J., Faro, S., Ferraria, P., Fitzgerald, M., Gege, D.,
 Galagan, J., Gardina, S., Graham, L., Grand-Pierre, N., Hafez, N.,
 Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
 Johnson, R., Jones, C., Kamat, A., Keibis, C., Landers, T.,
 Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Luis, A., Mabbott, R.,
 Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C.,
 McCarthy, M., Meldrum, J., Maneus, L., Mihova, T., Mlenga, V.,
 Murphy, T., Navljar, J., Nguyen, C., Nguyen, T., Nicoll, R., Norbu, C.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunckhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Rogov, P.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R.,
 Seaman, S., Severny, P., Smith, C., Spencer, B., Strange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talmas, J., Testyare, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.

TITLE	Direct Submission
JOURNAL	09-APR-2004 Whitehead Institute for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	4 (bases 1 to 24580)
AUTHORS	Birren, B., Nusbaum, C., Lander, E., Abouelhail, A., Alien, N., Anderson, M., Andersson, S., Arachchi, H. M., Barna, N., Bastien, V. J., Bloom, T., Boguskiy, L., Boukhefet, B., Camarata, J., Chang, J., Choepel, Y., Collimore, A., Cook, A., Cooke, P., Corum, B., DeAngelis, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, J. J., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardya, S., Graham, L., Grand-Pierre, N., Hafez, N., Hegopian, D., Hedges, B., Hall, J., Horton, L., Huie, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Keilis, C., Landers, T. V., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R. R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C. C., McCarthy, M., Meldrum, J., Minor, J., Mnihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rice, C., Rogov, P., Roman, J., Schauer, S. S., Schupback, R. R., Seaman, S., Severy, P., Smith, C., Spencer, B., Strange, Theodore, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliiev, H., Venkataraman, V. S., Viel, R. R., Wilson, B. J., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, M., and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (03-JUN-2004) Whitehead Institute/MIT Center
RESEARCH CENTER Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jun 3, 2004 this sequence version replaced gi:46309
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Contact: Whitehead Institute/MIT Center for Genome
Research, Whitehead Institute, One Cambridge Center, Cambridge, MA 02142, USA

```

Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information -----
Center project name: L18591
Center clone name: 246_F_14
----- FEATURES -----
source
  1. . 214580
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="Itaxon:10090"
    /chromosome="12"
    /map="12"
    /clone_id="RP23-246F14"
    /clone_lname="RPCI-23 Female Mouse BAC"
    complement("1..217B")
    /rpt_family="L1"
    misc_feature
      1. . 6
        /note="clone boundary
clone end SP6
site:EcORI"
        complement("2147..2719")
        /rpt_family="L1"
        2841..2933
        /rpt_family="M6R73"
        3052..3086
        /rpt_family="AT rich"
        compTlement("3966..4048")
        /rpt_family="T9A"
        4066..4092
        /rpt_family="(TTAA)n"
        4093..4115
        /rpt_family="((TTCC)n"
        repeat_region
        compTlement("4181..4330")
        /rpt_family="B3A"
        6284..6339
        /rpt_family="AT rich"
        7268..7417
        /rpt_family="LIMDa"
        7580..7600
        /rpt_family="AT rich"
        7650..7691
        /rpt_family="(TCTG)n"
        7696..7725
        /rpt_family="((TCCTCG)n"
        repeat_region
        compTlement("7756..8044")
        /rpt_family="B4"
        8402..8681
        /rpt_family="MMT"
        compTlement("9408..9432")
        /rpt_family="B1F"
        9433..9833
        /rpt_family="WTC"
        complement("9834..9922")
        /rpt_family="B1"
        12768..13055
        /rpt_family="Lx2"
        13768..13900
        /rpt_family="B1_MM"
        repeat_region
        13901..13961
        /rpt_family="(CRAA)n"
        repeat_region
        complement("15199..15315")
        /rpt_family="MLT1G3"
        16201..16394
        /rpt_family="(CGGTG)n"
        16505..16564
        /rpt_family="(CRAAA)n"
        repeat_region
        16610..16996
        /rpt_family="Lx6"
        17007..17055
        /rpt_family="(CRAAA)n"
        repeat_region
        17057..17505
        /rpt_family="(CRAAA)n"

```

repeat_region	/rpt_family="Lx6" complement(17621..17854)	O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Riley,R., Rize,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spender,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
repeat_region	17950..17993	Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., Deurrellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand,Pierre,N., Halez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbott,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Menuez,L., Mihowa,V., Murphy,T., Naylor,J., Nguyen,C., Nicols,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Recta,R., Rize,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Testayse,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
repeat_region	18085..18152	Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
repeat_region	18181..18210	3 (bases 1 to 245862)
repeat_region	18344..18374	Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
repeat_region	18925..19138	JOURNAL
repeat_region	19135..19166	REFERENCE
repeat_region	19167..19364	AUTHORS
repeat_region	/rpt_family="RLTR1OC"	Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., Deurrellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand,Pierre,N., Halez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbott,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Menuez,L., Mihowa,V., Murphy,T., Naylor,J., Nguyen,C., Nicols,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Recta,R., Rize,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Testayse,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
repeat_region	/rpt_complement(20397..20500)	JOURNAL
repeat_region	20397..20859	COMMENT
repeat_region	/rpt_family="Trigger"	Submitted (24-FEB-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
repeat_region	/rpt_complement(20569..20859)	3 (bases 1 to 2857041).
repeat_region	/rpt_family="Trigger7"	JOURNAL
repeat_region	/rpt_complement(20538..21158)	COMMENT
repeat_region	/rpt_family="34A"	3 (bases 1 to 2857041).
repeat_region	21360..21382	COMMENT
repeat_region	/rpt_family="AT_rich"	Submitted (24-FEB-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
repeat_region	/rpt_complement(21486..21826)	3 (bases 1 to 2857041).
repeat_region	/rpt_family="Lx8"	JOURNAL
repeat_region	21859..21891	COMMENT
repeat_region	/rpt_family="(TG)n"	3 (bases 1 to 2857041).
repeat_region	/rpt_complement(21504..22282)	COMMENT
repeat_region	/rpt_family="ORR1B2"	Submitted (24-FEB-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
repeat_region	/rpt_complement(22283..22499)	3 (bases 1 to 2857041).
Query Match	71.1%	TITLE
Best Local Similarity	87.5%	JOURNAL
Matches 21;	Score 19.2;	Length 214580;
Conservative 0;	Pred. No. 2.6e-02;	HTG 24-FBB-2004
Mismatches 3;	Indels 0;	SEQUENCING
Qy 3 GGCCTCAATGAAATTGATATGATTC 26	Gaps 0;	IN PROGRESS ***; 7 unordered pieces.
Db 8449 GGCCTCAATGAAATTGATATGATTC 8472	;	AC123598 AC123598 Mus musculus chromosome 12 clone RP24-501ES map 12, ***
RESULT 4		AC123598 AC123598 linear
ACCESSION		AC123598 linear
VERSION		GI:42761736
KEYWORDS		HTG, HTGS, PHASE1, HTGS_FULLTOP, HTGS_ACTIVEFIN.
SOURCE		Mus musculus (house mouse)
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 245862)	JOURNAL
AUTHORS	Birren,B., Nusbaum,C. and Lander,E.	REFERENCE
TITLE	Unpublished	AC123598
JOURNAL	2 (bases 1 to 245862)	AC123598
REFERENCE	Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Deurrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Gande,S., Gord,S., Govertte,M., Graham,L., Grand,Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G., MacIe,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Menuez,L., Mihowa,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,	
AUTHORS		FEATURES
source		1..245862 /organism="Mus musculus" /mol_type="genomic DNA"

/db_xref="taxon:10090" /chromosome="12" /map="12" /clone="RPCI-24 501ES" /clone_lib="RPCI-24 Male Mouse BAC"	ORIGIN	Query Match Best Local Similarity 71.1%; Score 19.2; DB 2; Length 245862; Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	FEATURES source 1. .663	REFERENCE AUTHORS Morris,D.W. and Engelhard,E.K. TITLE Novel compositions and methods for cancer JOURNAL WO 0308553-A 15.10.30-JAN-2003; Sagres Discovery (US) Location/Qualifiers
Qy 3 GCGCCATGAAAGTATGATTC 26 Db 204043 GCCTCAATTGAGATATGATTC 204066	RESULT 5 HUMCD282 LOCUS HUMCD282 DEFINITION Human T-cell membrane glycoprotein CD28 mRNA, exon 2. VERSION M37813 KEYWORDS glycoprotein CD28. SEGMENT 2 of 4 ORGANISM Homo sapiens (human)	Query Match Best Local Similarity 87.5%; Score 19.2; DB 2; Length 542 bp; Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	FEATURES source 1. .663	REFERENCE AUTHORS Gan,S.U., Hare,J., Krivoshechapov,L., Hui,K.M., Galea-Lauri,J., Parzaneh,F. and Darling,D. TITLE New human CD28 isoforms Generated by a novel splicing event in the 5'UTR COMMENT Original source text: Human DNA. FEATURES source
Qy 5 HUMCD282 LOCUS HUMCD282 DEFINITION Human T-cell membrane glycoprotein CD28 mRNA, exon 2. VERSION M37813 KEYWORDS glycoprotein CD28. SEGMENT 1 of 4 ORGANISM Homo sapiens (human)	RESULT 7 HSA517504 LOCUS HSA517504 DEFINITION Homo sapiens mRNA for T-cell-specific surface glycoprotein (CD28 gene), splice variant 1. VERSION AJ517504.1 KEYWORDS alternative splicing; CD28 gene; T-cell-specific surface glycoprotein. ORGANISM Homo sapiens (human)	Query Match Best Local Similarity 100.0%; Score 19; DB 6; Length 663; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	FEATURES source 1. .705	REFERENCE AUTHORS Gan,S.U., Hare,J., Krivoshechapov,L., Hui,K.M., Galea-Lauri,J., Parzaneh,F. and Darling,D. TITLE New human CD28 isoforms Generated by a novel splicing event in the 5'UTR COMMENT Original source text: Human DNA. FEATURES source 1. .705
Qy 6 AX695883 LOCUS AX695883 DEFINITION Sequence 1510 from Patent WO03008583. VERSION AX695883.1 KEYWORDS Homo sapiens (human)	RESULT 6 AX695883 LOCUS AX695883 DEFINITION Sequence 1510 from Patent WO03008583. VERSION AX695883.1 KEYWORDS Homo sapiens (human)	Query Match Best Local Similarity 100.0%; Score 19; DB 9; Length 705; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	FEATURES source 1. .705	REFERENCE AUTHORS Morris,D.W. and Engelhard,E.K. TITLE Novel compositions and methods for cancer JOURNAL WO 0308553-A 15.10.30-JAN-2003; Sagres Discovery (US) Location/Qualifiers
Qy 7 AX695883 LOCUS AX695883 DEFINITION Sequence 1510 from Patent WO03008583. VERSION AX695883.1 KEYWORDS Homo sapiens (human)	RESULT 7 AX695883 LOCUS AX695883 DEFINITION Sequence 1510 from Patent WO03008583. VERSION AX695883.1 KEYWORDS Homo sapiens (human)	Query Match Best Local Similarity 100.0%; Score 19; DB 9; Length 705; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	FEATURES source 1. .705	REFERENCE AUTHORS Morris,D.W. and Engelhard,E.K. TITLE Novel compositions and methods for cancer JOURNAL WO 0308553-A 15.10.30-JAN-2003; Sagres Discovery (US) Location/Qualifiers
Qy 8 AX695883 LOCUS AX695883 DEFINITION Sequence 1510 from Patent WO03008583. VERSION AX695883.1 KEYWORDS Homo sapiens (human)	RESULT 8 AX695883 LOCUS AX695883 DEFINITION Sequence 1510 from Patent WO03008583. VERSION AX695883.1 KEYWORDS Homo sapiens (human)	Query Match Best Local Similarity 100.0%; Score 19; DB 9; Length 705; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	FEATURES source 1. .705	REFERENCE AUTHORS Morris,D.W. and Engelhard,E.K. TITLE Novel compositions and methods for cancer JOURNAL WO 0308553-A 15.10.30-JAN-2003; Sagres Discovery (US) Location/Qualifiers
Qy 9 AATTGAAGTTATGATTC 27 Db 351 AATTGAAGTTATGATTC 369	RESULT 9 AATTGAAGTTATGATTC 27 Db 339 AATTGAAGTTATGATTC 357	Query Match Best Local Similarity 100.0%; Score 19; DB 6; Length 663; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	FEATURES source 1. .705	REFERENCE AUTHORS Gan,S.U., Hare,J., Krivoshechapov,L., Hui,K.M., Galea-Lauri,J., Parzaneh,F. and Darling,D. TITLE New human CD28 isoforms Generated by a novel splicing event in the 5'UTR COMMENT Original source text: Human DNA. FEATURES source 1. .705

Db	381 AATTGAAGTTATGTATCCCT 399	Query Match Best Local Similarity 70.4%; Pred. No. 2.1e+02; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 8		
BD267803	BD267803 1514 bp DNA linear PAT 17-JUL-2003	DEFINITION Methods and products related to metabolic interactions in disease.
LOCUS		
DEFINITION		
ACCESSION		
VERSION	BD267803	
KEYWORD	BP267803.1 GI:32077571	
SOURCE	JP 200251203-A/6.	
ORGANISM	Hom sapiens (human)	
REFERENCE	Newell, M.K.	
AUTHORS		
TITLE	Methods and products related to metabolic interactions in disease	
JOURNAL	Patent: JP 200251203-A 6 23-APR-2002;	
COMMENT	OS Homo sapiens (human)	
DN	JP 200251203-A/6	
PD	23-APP-2002	
PF	30-MAR-1999 JP 20000544356	
BR	17-APR-1998 US 60/082250, 29-JUL-1998 US	
24-SEP-1998 US 60/101580		
PI	MARTHA K NEWELL	
PC	A61K3/395,A61K3/51.9,A61K31/55,A61K31/704,A61K31/7056,A61K31/7088,	
PC	A61K31/17,A61K38/21,A61K45/06,A61P3/10,A61P35/00,C12N15/00, PC	
GO:00N33/54 ¹		
EC	A61K3/12,A61K3/66,C12N15/00	
CC	Methods and products related to metabolic interactions in CC	
Disease		
PH	Key source	
FT	Location/Qualifiers	
FEATURES	1..1514 /organism='Homo sapiens (human)' .	
source	1..1514 /organism='Homo sapiens' /mol_type="genomic DNA" /db_xref="taxon:9606"	
ORIGIN		
RESULT 9		
107254	107254 1514 bp DNA linear PAT 02-DEC-1994	Query Match Best Local Similarity 100.0%; Pred. No. 2.1e+02; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
LOCUS		
DEFINITION	Sequence 11 from Patent EP 0330191.	
ACCESSION	107254	
VERSION	107254.1 GI:590064	
KEYWORD	Unclassified.	
ORGANISM	Unknown.	
REFERENCE	Seed, B., D. Do, M. B., Allen, J., Aruffo, A., Camerini, D., Lauffer, L. D., Oquendo, C. P., Simmons, D., Stamenkovic, I., and Stengelin, S. D.	
AUTHORS		
TITLE	Rapid immunoselection cloning method	
JOURNAL	Patent: EP 0330191-A2 11 30-AUG-1989;	
FEATURES	Location/Qualifiers	
source	1..1514 /organism="unknown" /mol_type="unassigned DNA"	
ORIGIN		

ORIGIN /db_xref="taxon:9606"

RESULT 13
AX557214

Query Match Similarity 70.4%; Score 19; DB 6; Length 1514;
Best Local Matches 19; Conservative 0; Pred. No. 2.1e+02; Mismatches 0;
Indels 0; Gaps 0; Version AX557214.1 GI:25900195

LOCUS AX557214 3803 bp DNA
DEFINITION Sequence 1 from Patent WO02066059.
ACCESSION AX557214
VERSION AX557214.1
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS O'Hara, R.M. and Nagel, A.M.
TITLE Agents that specifically block cd38-mediated signaling and uses
therefor
JOURNAL Patent: WO 02066059-A 1 29-AUG-2002;
Genetics Institute, LLC (US)
FEATURES Location/Qualifiers
source 1..3803
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN Query Match Score 19; DB 6; Length 3803;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;

FEATURES source 9 AATTGAACTTATGTATCC 27
Db 560 AATTGAACTTATGTATCC 578

RESULT 14
AX695882

Query Match Score 19; DB 6; Length 3803;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;

LOCUS AX695882 3804 bp DNA
DEFINITION Sequence 1509 from Patent WO03008583.
ACCESSION AX695882
VERSION AX695882.1 GI:29419047

KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Morris, D.W. and Engelhard, E.K.
TITLE Novel compositions and methods for cancer
JOURNAL Patent: WO 03008583-A 1509 30-JAN-2003;
Sagres Discovery (US)
FEATURES Location/Qualifiers
source 1..3804
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN Query Match Score 19; DB 6; Length 3804;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;

FEATURES source 9 AATTGAACTTATGTATCC 27
Db 561 AATTGAACTTATGTATCC 579

RESULT 15
CQ715869

Query Match Similarity 70.4%; Score 19; DB 9; Length 1514;
Best Local Matches 19; Conservative 0; Pred. No. 2.1e+02; Mismatches 0;
Indels 0; Gaps 0; Version CQ715869.1 GI:42276726

LOCUS CQ715869 3806 bp DNA
DEFINITION Sequence 1803 from Patent WO02068579.
ACCESSION CQ715869
VERSION CQ715869.1 GI:42276726

KEYWORDS

SOURCE	Homo sapiens (human)		TITLE	Novel compositions and methods for cancer
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		JOURNAL	WO 0300853-A 1508 30-JAN-2003;
REFERENCE	1.		FEATURES	sgres discovery (US)
AUTHORS	Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.		source	Location/Qualifiers
TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof			1. .51365
JOURNAL	Patent: WO 02028579-A 1803 06-SEP-2002;			/organism="Homo sapiens"
FEATURES	PE Corporation (NY) (US)			/mol type="unassigned DNA"
SOURCE	Location/Qualifiers			/db_xref="taxon:9606"
	1. .3806			
	/organism="Homo sapiens"			
	/mol type="unassigned DNA"			
	/db_xref="taxon:9606"			
ORIGIN				
Query Match	70.4%; Score 19; DB 6; Length 3806;		RESULT 18	
Best Local Similarity	100.0%; Pred. No. 2.3e+02;		AF411057	
Matches	19; Conservative 0; Mismatches 0;		LOCUS	66887 bp
	Indels 0; Gaps 0;		DEFINITION	Linear
Qy	9 AATTGAAAGTTATGTATCCCT 27		Hom sapiens clone BAC 22606 CD28 antigen	PRI 13-DEC-2001
Db	561. AATTGAAAGTTATGTATCCCT 579		VERSION	(CD28 gene, exons 1, 2 and 3, and partial cds.
			VERSION	AF411057.1 GI:17646225
			KEYWORDS	
			source	Homo sapiens (human)
			ORGANISM	Homo sapiens
			REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
			AUTHORS	1. (bases 1 to 66887) Ling,V., Wu,P.W., Pinnerty,H.F., Agostino,M.J., Graham,J.R., Chen,S., Jussiff,J.J., Fisk,C.P. and Collins,M.
			VERSION	Title Assembly and Annotation of Human Chromosome 2q33 Sequence Containing the CD28, CMLA, and ICOS Gene Cluster: Analysis by Computational, Comparative, and Microarray Approaches
			JOURNAL	Genomics 78 (3), 155-168 (2001)
			PUBLMED	11752222
			REFERENCE	2. (bases 1 to 66887) Ling,V., Wu,P.W., Pinnerty,H.F., Agostino,M.J., Graham,J.R., Chen,S., Jussiff,J.J., Fisk,C.P. and Collins,M.
			AUTHORS	Title Direct Submission
			VERSION	Journal Submitted (15-AUG-2001) Department of Immunology, Genetics Institute Wyeth Research, 200 Cambridge Park Drive, Cambridge, MA 02140, USA
			JOURNAL	02140, USA
			FEATURES	Location/Qualifiers
			source	1. .66887
				/organism="Homo sapiens"
				/mol type="genomic DNA"
				/db_xref="taxon:9606"
				/chromosome="2"
				/map="2q33"
				/clone="BAC 22606"
				7838 .8329_
				/note="similar to NADH:ubiquinone oxidoreductase"
				gene 42348 .>66887
				/gene="CD28"
				/join=42348 .42621 .62505 . .62861 .65540 . .>65664)
				mRNA 42348 . .42621 .62505 . .62861 .65540 . .>65664)
RESULT 17				
AX65881	AX65881 AX695881 51365 bp			
LOCUS	DNA			
DEFINITION	Sequence 1508 from Patent WO03008583.			
VERSION	AX65881 AX695881.1 GI:29419046			
KEYWORDS				
SOURCE				
ORGANISM	Homo sapiens (human)			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1. Morris,D.W. and Engelhard,E.K.			
AUTHORS				

NLSREFASLHKGLDSAVECVVYGNYSQOLQYSKTKGENGDGRLGNESVTFTYQLQNLYNQCDIYFCKLEVMYPPVLDNEKSNGTIIHVKGKHLCESPPLFGPSRKFPMWLVVVG
GVLAFCYSLLVTAFLIFW"

62505 . 62611
/gene="CD28"
/number=2
65540 . 65664
/gene="CD28"
/number=3

ORIGIN

Query Match 70.4%; Score 19; DB 9; Length 66887;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATTGAAGTTATGATATCCCT 27
Db 62791 AATTGAAGTTATGATATCCCT 62809

RESULT 19

AC020676 88624 bp DNA linear HTG 13-JUL-2000
DEFINITION Homo sapiens clone RP11-8P23, LOW-PASS SEQUENCE SAMPLING.
AC020676 GI:9123792
VERSION 1
KEYWORDS HTG; HTGS; PHASED.
SOURCE Birren, B.; Linton, L.; Nusbaum, C. and Lander, E.
ORGANISM Homo sapiens (human)
Homosapiens, clone RP11-8P23

REFERENCE 1 (bases 1 to 88624)

AUTHORS Birren, B.; Linton, L.; Nusbaum, C.; Landier, B.; Abraham, H.; Allen, N.,
Anderson, S., Baldwin, J.; Barna, N.; Beckerly, R.; Beda, F.;
Boguslavskiy, L.; Boukhalter, B.; Brown, A.; Burkett, G.; Castle, A.,
Choeper, Y.; Collangeli, M.; Collins, S.; Collymore, A.; Cooke, P.,
DeAraujo, X.; Dewar, K.; Domingo, M.; Doyle, M.; Fenestor, J.,
Ferreira, P.; FitzHugh, W.; Forrest, C.; Gage, D.; Galagan, J.,
Gardyna, S.; Grant, G.; Hados, B.; Hastings, A.; Horton, L.,
Howland, J. C.; Johnson, R.; Jones, C.; Kahn, L.; Karatas, A.; Klein, J.,
Landers, T.; Lehoczky, J.; Levine, R.; Lieu, C.; Liu, G.; Locke, K.,
MacDonald, P.; Marquis, N.; McEwan, P.; McKernan, K.,
McPhreeters, R.; Meldrim, J.; Menus, L.; Morrow, J.; Naylor, J.,
Norman, C. H.; O'Connor, T.; O'Donnell, P.; Oliver, T. M.; Peterson, K.,
Pierre, N.; Pisani, C.; Pollara, V.; Raymond, C.; Riley, R.; Rothman, D.,
Roy, A.; Santos, R.; Severy, P.; Spencer, B.; Strange-Thomann, N.,
Stojanovic, N.; Subramanian, A.; Talamas, J.; Tesfaye, S.; Theodore, J.,
Tirrell, A.; Vasquez, H.; Vier, R.; Vo, A.; Wu, X.; Wyman, D.; Ye, W. J.,
Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (08-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jul 13, 2000 this sequence version replaced gi:6682672.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute / MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2995
Center clone name: 8_P_23

* arbitrary. Low-pass sequence sampling is useful for identifying clones that may be gene-rich and allows overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone will be sequenced to completion. In the event that the record is updated, the accession number will be preserved.
* * 1 882: contig of 882 bp in length
* * * 883 982: gap of 100 bp
* * * 983 1886: contig of 904 bp in length
* * * 1887 1986: gap of 100 bp
* * * 1987 2905: contig of 919 bp in length
* * * 2906 3001: gap of 100 bp
* * * 3006 3903: contig of 898 bp in length
* * * 3904 4003: gap of 100 bp
* * * 4004 4925: contig of 922 bp in length
* * * 4926 5025: gap of 100 bp
* * * 5924: contig of 899 bp in length
* * * 5925 6024: gap of 100 bp
* * * 6025 6937: contig of 913 bp in length
* * * 6938 7037: gap of 100 bp
* * * 7038 7949: contig of 912 bp in length
* * * 7950 8049: gap of 100 bp
* * * 8050 8920: contig of 871 bp in length
* * * 8921 9020: gap of 100 bp
* * * 9021 9894: contig of 874 bp in length
* * * 9995 10908: gap of 100 bp
* * * 10909 11008: contig of 914 bp in length
* * * 11009 11921: gap of 100 bp
* * * 11922 12021: contig of 913 bp in length
* * * 12022 12931: contig of 910 bp in length
* * * 12932 13031: gap of 100 bp
* * * 13032 13947: contig of 916 bp in length
* * * 13948 14048: gap of 100 bp
* * * 14048 14949: contig of 902 bp in length
* * * 15050 15049: gap of 100 bp
* * * 15961 16060: gap of 100 bp
* * * 16061 16971: contig of 911 bp in length
* * * 16972 17071: gap of 100 bp
* * * 17072 17996: contig of 925 bp in length
* * * 17997 18096: gap of 100 bp
* * * 18097 18986: contig of 890 bp in length
* * * 18987 19086: gap of 100 bp
* * * 19087 19942: contig of 906 bp in length
* * * 19943 20093: gap of 100 bp
* * * 20093 20958: contig of 866 bp in length
* * * 20959 21058: gap of 100 bp
* * * 21059 21948: contig of 890 bp in length
* * * 21949 22048: gap of 100 bp
* * * 22049 22957: contig of 903 bp in length
* * * 22958 23057: gap of 100 bp
* * * 23058 23967: contig of 910 bp in length
* * * 23968 24067: gap of 100 bp
* * * 24068 25001: contig of 934 bp in length
* * * 25002 25101: gap of 100 bp
* * * 25102 26000: contig of 893 bp in length
* * * 26001 26100: gap of 100 bp
* * * 26101 27013: contig of 913 bp in length
* * * 27014 27113: gap of 100 bp
* * * 27114 28031: contig of 918 bp in length
* * * 28032 28131: gap of 100 bp
* * * 28132 29041: contig of 910 bp in length
* * * 29042 29141: gap of 100 bp
* * * 29142 30038: contig of 897 bp in length
* * * 30039 30138: gap of 100 bp
* * * 30139 31031: contig of 893 bp in length
* * * 31032 32049: gap of 100 bp
* * * 31132 32149: gap of 100 bp
* * * 32050 32150: contig of 915 bp in length
* * * 33064: gap of 100 bp
* * * 33164: gap of 100 bp

* NOTE: This record contains 88 individual sequencing reads that have not been assembled into contigs. Runs of N are used to separate the reads and the order in which they appear is completely

1	882:	contig of 882 bp in length	gap of 100 bp in length
**	883:	gap of 100 bp	contig of 916 bp in length
**	983:	contig of 904 bp in length	gap of 100 bp
**	1886:	gap of 100 bp	contig of 905 bp in length
**	1887:	contig of 919 bp in length	gap of 100 bp
**	1987:	gap of 100 bp	contig of 908 bp in length
**	2905:	contig of 919 bp in length	gap of 100 bp
**	3005:	gap of 100 bp	contig of 904 bp in length
**	3006:	contig of 898 bp in length	gap of 100 bp
**	3903:	gap of 100 bp	contig of 900 bp in length
**	4003:	contig of 922 bp in length	gap of 100 bp
**	4925:	contig of 922 bp in length	contig of 904 bp in length
**	5025:	gap of 100 bp	contig of 915 bp in length
**	5924:	contig of 899 bp in length	gap of 100 bp
**	6024:	gap of 100 bp	contig of 899 bp in length
**	6025:	contig of 913 bp in length	gap of 100 bp
**	6938:	gap of 100 bp	contig of 922 bp in length
**	7038:	contig of 912 bp in length	gap of 100 bp
**	7950:	gap of 100 bp	contig of 920 bp in length
**	8050:	contig of 871 bp in length	gap of 100 bp
**	8920:	gap of 100 bp	contig of 990 bp in length
**	9020:	gap of 100 bp	contig of 990 bp in length
**	9894:	contig of 874 bp in length	gap of 100 bp
**	9995:	gap of 100 bp	contig of 899 bp in length
**	9999:	contig of 914 bp in length	gap of 100 bp
**	10908:	gap of 100 bp	contig of 925 bp in length
**	11008:	gap of 100 bp	contig of 925 bp in length
**	11009:	contig of 913 bp in length	gap of 100 bp
**	11922:	contig of 910 bp in length	gap of 100 bp
**	12022:	gap of 100 bp	contig of 933 bp in length
**	12931:	contig of 910 bp in length	gap of 100 bp
**	13031:	gap of 100 bp	contig of 926 bp in length
**	13032:	contig of 916 bp in length	gap of 100 bp
**	13948:	gap of 100 bp	contig of 898 bp in length
**	14047:	contig of 902 bp in length	gap of 100 bp
**	14949:	contig of 100 bp	contig of 914 bp in length
**	15049:	gap of 100 bp	contig of 917 bp in length
**	15050:	contig of 911 bp in length	gap of 100 bp
**	15961:	gap of 100 bp	contig of 908 bp in length
**	16061:	contig of 911 bp in length	gap of 100 bp
**	16971:	gap of 100 bp	contig of 906 bp in length
**	17071:	contig of 925 bp in length	gap of 100 bp
**	17996:	contig of 902 bp in length	gap of 100 bp
**	17997:	gap of 100 bp	contig of 922 bp in length
**	18097:	contig of 890 bp in length	gap of 100 bp
**	18987:	gap of 100 bp	contig of 916 bp in length
**	19087:	contig of 906 bp in length	gap of 100 bp
**	19992:	contig of 906 bp in length	contig of 922 bp in length
**	20092:	gap of 100 bp	contig of 914 bp in length
**	20958:	contig of 866 bp in length	gap of 100 bp
**	20993:	gap of 100 bp	contig of 917 bp in length
**	21059:	contig of 890 bp in length	gap of 100 bp
**	21948:	contig of 934 bp in length	gap of 100 bp
**	22048:	gap of 100 bp	contig of 916 bp in length
**	22957:	contig of 909 bp in length	gap of 100 bp
**	23057:	gap of 100 bp	contig of 916 bp in length
**	23967:	contig of 910 bp in length	gap of 100 bp
**	23968:	gap of 100 bp	contig of 917 bp in length
**	24067:	contig of 934 bp in length	gap of 100 bp
**	25002:	gap of 100 bp	contig of 902 bp in length
**	25102:	contig of 899 bp in length	gap of 100 bp
**	26001:	gap of 100 bp	contig of 894 bp in length
**	26101:	contig of 913 bp in length	gap of 100 bp
**	27013:	contig of 913 bp in length	gap of 100 bp
**	27112:	gap of 100 bp	contig of 933 bp in length
**	27114:	contig of 918 bp in length	gap of 100 bp
**	28031:	gap of 100 bp	contig of 896 bp in length
**	28131:	contig of 910 bp in length	gap of 100 bp
**	29041:	gap of 100 bp	contig of 916 bp in length
**	29141:	contig of 917 bp in length	gap of 100 bp
**	30038:	contig of 913 bp in length	gap of 100 bp
**	30138:	gap of 100 bp	contig of 929 bp in length
**	31039:	contig of 893 bp in length	gap of 100 bp
**	31131:	gap of 100 bp	contig of 902 bp in length
**	32049:	contig of 918 bp in length	gap of 100 bp
**	32149:	gap of 100 bp	contig of 916 bp in length
**	32250:	contig of 915 bp in length	gap of 100 bp
**	32350:	gap of 100 bp	contig of 913 bp in length
**	33165:	contig of 906 bp in length	gap of 100 bp
**	34070:	gap of 100 bp	contig of 912 bp in length
**	34171:	contig of 924 bp in length	gap of 100 bp
**	35095:	gap of 100 bp	contig of 932 bp in length
**	35195:	contig of 929 bp in length	gap of 100 bp
**	36226:	gap of 100 bp	contig of 929 bp in length
**	37155:	contig of 929 bp in length	gap of 100 bp
**	37255:	gap of 100 bp in length	Score 19; DB 2; Length 8
**	37256:	contig of 916 bp in length	Best Local Similarity 100.0%; Pre-Matches 0;
**	38171:	gap of 100 bp	Conservative Matches 19;
**	38211:	contig of 905 bp in length	Mismatches 0; Index 3.1e+02;
**	38272:	gap of 100 bp	
**	39177:	contig of 908 bp in length	
**	39277:	contig of 908 bp in length	
**	40184:	gap of 100 bp	
**	40285:	contig of 900 bp in length	
**	41184:	gap of 100 bp	
**	41185:	contig of 904 bp in length	
**	41285:	gap of 100 bp	
**	42288:	contig of 908 bp in length	
**	42377:	gap of 100 bp	
**	42389:	contig of 899 bp in length	
**	43178:	gap of 100 bp	
**	43278:	contig of 922 bp in length	
**	44199:	gap of 100 bp	
**	44200:	contig of 920 bp in length	
**	44300:	gap of 100 bp	
**	45220:	contig of 100 bp	
**	45229:	gap of 100 bp	
**	46210:	contig of 990 bp in length	
**	46709:	gap of 100 bp	
**	46310:	contig of 899 bp in length	
**	47208:	gap of 100 bp	
**	47308:	contig of 925 bp in length	
**	48233:	gap of 100 bp	
**	48234:	contig of 933 bp in length	
**	49267:	gap of 100 bp	
**	49366:	contig of 926 bp in length	
**	50232:	gap of 100 bp	
**	50293:	contig of 908 bp in length	
**	50393:	gap of 100 bp	
**	51290:	contig of 925 bp in length	
**	51390:	gap of 100 bp	
**	51291:	contig of 932 bp in length	
**	52279:	gap of 100 bp	
**	52319:	contig of 926 bp in length	
**	52380:	gap of 100 bp	
**	53275:	contig of 906 bp in length	
**	53276:	gap of 100 bp	
**	53375:	contig of 914 bp in length	
**	54290:	gap of 100 bp	
**	54291:	contig of 917 bp in length	
**	54311:	gap of 100 bp	
**	55312:	contig of 908 bp in length	
**	56312:	gap of 100 bp	
**	56320:	contig of 906 bp in length	
**	56420:	gap of 100 bp	
**	57325:	contig of 906 bp in length	
**	57326:	gap of 100 bp	
**	57426:	contig of 917 bp in length	
**	58426:	gap of 100 bp	
**	58433:	contig of 922 bp in length	
**	58442:	gap of 100 bp	
**	58443:	contig of 916 bp in length	
**	58443:	gap of 100 bp	
**	59358:	contig of 908 bp in length	
**	59458:	gap of 100 bp	
**	60480:	contig of 906 bp in length	
**	60381:	gap of 100 bp	
**	61381:	contig of 906 bp in length	
**	61387:	gap of 100 bp	
**	62265:	contig of 879 bp in length	
**	62336:	gap of 100 bp	
**	62465:	contig of 922 bp in length	
**	63339:	gap of 100 bp	
**	63360:	contig of 916 bp in length	
**	63460:	gap of 100 bp	
**	64152:	contig of 933 bp in length	
**	64353:	gap of 100 bp	
**	64453:	contig of 902 bp in length	
**	65354:	gap of 100 bp	
**	65355:	contig of 904 bp in length	
**	66310:	gap of 100 bp	
**	66371:	contig of 929 bp in length	
**	66471:	gap of 100 bp	
**	67499:	contig of 929 bp in length	
**	67500:	gap of 100 bp	
**	68335:	contig of 896 bp in length	
**	68395:	gap of 100 bp	
**	68496:	contig of 896 bp in length	
**	69392:	gap of 100 bp	
**	70404:	contig of 913 bp in length	
**	70544:	gap of 100 bp	
**	71416:	contig of 912 bp in length	
**	71517:	gap of 100 bp	

Query Match	Best Local Similarity	Score 19;	DB 2;	Length 88624;	
Best Local Matches	100 %;	Pred. No. 3.1e+02;	Misaligned 0;	Indels 0;	Gaps 0;
Conservative Matches	19;	Score 19;	DB 2;	Length 88624;	

	9	AATTGAGGTATGATCTT	27					
y	AF225899	106539 bp	DNA	linear	PRI	23-FEB-2000		
b	OCUS	Homo sapiens PAC clone 219d7, complete sequence.						
	REPINTION	Homo sapiens						
	ACCESSION	AF225899						
	VERSION	AF225899.1	GI:7021528					
RESULT 21	KEYWORDS	Homo sapiens (human)						
	SOURCE	Homo sapiens						
	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
	REFERENCE	1 (bases 1 to 106539)						
	AUTHORS	Fitzpatrick, E.S., Hammond, H.A., DeAngelis, D.M., Soderman, A.R., Wright, J.L., Liu, X., Larson, D., McGowan, J., Ziegler, S., Pritchard, L., Hess, J.F., Todd, J.J., Caskey, C.T. and Matzker, M.L.						
	TITLE	Direct Submission						
	JOURNAL	Submitted (18-JAN-2000) Department of Human Genetics, Merck & Co., Inc., SunneyTown Pike, West Point, PA 19486, USA						
	FEATURES	Location/Qualifiers						
		1..106539						
		/organism="Homo Sapiens"						
		/mol type="genomic DNA"						
		/db_xref="taxon:9606"						
		/clone="PAC 219d7"						
	repeat_region	repeat_region	(34 .. 104)					
		/rpt_family="MRS5"						
		complement(384 .. 662)						
	repeat_region	repeat_region	1021..11304					
		/rpt_family="Alu"						
		complement(1021 .. 11304)						
	repeat_region	repeat_region	1800..2135					
		/rpt_family="Alu"						
		complement(1800 .. 2135)						
	repeat_region	repeat_region	2139..2439					
		/rpt_family="Alu"						
		complement(2139 .. 2439)						
	repeat_region	repeat_region	3214..3294					
		/rpt_family="Alu"						
		complement(3214 .. 3294)						
	repeat_region	repeat_region	3588..3657					
		/rpt_family="MTR2B2"						
		complement(3588 .. 3657)						
	repeat_region	repeat_region	3982..4274					
		/rpt_family="Alu"						
		complement(3982 .. 4274)						
	repeat_region	repeat_region	5406..5550					
		/rpt_family="MER3"						
		complement(5406 .. 5550)						
	repeat_region	repeat_region	5444..5509					
		/rpt_family="MTR2B2"						
		complement(5444 .. 5509)						
	repeat_region	repeat_region	5558..5583					
		/rpt_family="Alu"						
		complement(5558 .. 5583)						
	repeat_region	repeat_region	5898..6046					
		/rpt_family="MER3"						
		complement(5898 .. 6046)						
	repeat_region	repeat_region	6534..6827					
		/rpt_family="Alu"						
		complement(6534 .. 6827)						
	repeat_region	repeat_region	8239..8568					
		/rpt_family="LTR12"						
		complement(8239 .. 8568)						
	repeat_region	repeat_region	11864..1259					
		/rpt_family="Alu"						
		complement(11864 .. 1259)						
	repeat_region	repeat_region	12838..13137					
		/rpt_family="Alu"						
		complement(12838 .. 13137)						
	repeat_region	repeat_region	14450..14691					
		/rpt_family="Alu"						
		complement(14450 .. 14691)						
	repeat_region	repeat_region	15004..15291					
		/rpt_family="Alu"						
		complement(15004 .. 15291)						
	repeat_region	repeat_region	16443..16422					
		/rpt_family="Alu"						
		complement(16443 .. 16422)						
	repeat_region	repeat_region	16813..17096					
		/rpt_family="Alu"						
		complement(16813 .. 17096)						
	repeat_region	repeat_region	17528..17706					
		/rpt_family="MIR"						
		complement(17528 .. 17706)						
	repeat_region	repeat_region	21861..21890					
		/rpt_family="Alu"						
		complement(21861 .. 21890)						

repeat_region / rpt family="Alu" 82595 .82857
repeat_region / rpt family="Alu" complement(84782 .85059)
repeat_region / rpt family="Alu"
repeat_region / rpt family="Alu"
repeat_region / rpt family="Alu"
repeat_region / rpt family="Alu"
repeat_region / rpt family="Alu" 87328 .87177
repeat_region / rpt family="Alu"
repeat_region / rpt family="Alu" complement(84025 .88300)
repeat_region / rpt family="Alu"
repeat_region / rpt family="Alu" complement(88382 .88537)
repeat_region / rpt family="Alu"
repeat_region / rpt family="L1" complement(89103 .92131)
repeat_region / rpt family="L1" complement(89607 .89870)
repeat_region / rpt family="Alu"
repeat_region / rpt family="Alu" complement(92183 .92468)
repeat_region / rpt family="Alu"
repeat_region / rpt family="L1" complement(92638 .92723)
repeat_region / rpt family="L1" complement(93254 .93542)
repeat_region / rpt family="Alu"
repeat_region / rpt family="MER43"
repeat_region / rpt family="MER43" 94819 .94382
repeat_region / rpt family="MER43" 97105 .97472
repeat_region / rpt family="THB1"
STS 99270 .99669
note="G08015 human STS CHLC.GGAA19E07.P17317 CLONE
GGAA19E07"
repeat_region complement(101199 .101438)
repeat_region / rpt family="Alu" 101753 .102038
repeat_region / rpt family="Alu"
repeat_region complement(102216 .102638)
repeat_region / rpt family="LOR1"
repeat_region complement(103792 .104101)
repeat_region / rpt family="L1"
repeat_region complement(104170 .104465)
repeat_region / rpt family="Alu" 104528 .104642
repeat_region / rpt family="L1" 105260 .105545
repeat_region / rpt family="Alu"
repeat_region / rpt family="Alu" 105891 .106161
repeat_region / rpt family="Alu"
ORIGIN
Query Match 70.4% Score 19; DB 9; Length 106539;
Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;
KEYWORDS
Homo sapiens (human)
ORGANISM
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108100)

AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11 , 1097-1108 (1998)
MEDLINE 99053792
PUBMED 9847074
REFERENCE 2 (bases 1 to 108100)
AUTHORS Levy,A., Kozlowski,A. and Spalding,L.
TITLE The sequence of Homo sapiens BAC clone RP11-638D14
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 108100)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (29-NOV-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 108100)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (23-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Mar 23, 2002 this sequence version replaced gi:18030160.
Center: Washington University Genome Sequencing Center
Center code: WIGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NHG0638D14

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.
This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subsection; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Cataneo,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.reagen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-789A18, 2000 bp overlap; the clone sequenced to the right is RP11-515K12, 2000 bp overlap. Actual start of this clone is at base position 146042 of RP11-789A18; actual end is at base position 23686 of RP11-515K12. Polymorphisms have been identified between AC093113 and AC103881. A transposon was identified in a growth of the clone RP11-638D14, which is not a part of the submitted sequence.

FEATURES
source
Location/Qualifiers
1. .108100
/organism="Homo sapiens"
/mol_type="genomic DNA"

```

/db_xref="taxon:9606"
/chromosome="2"
/map=12
/clone.lib="RPCI-11"
repeat_region 411. .867
/rept.family="MALR"
repeat_region 1022. 1329
/rept.family="Alu"
repeat_region 3858-.4066
/rept.family="MIR"
repeat_region 4074. .4179
/rept.family="MIR"
repeat_region 4527. .4695
/rept.family="MIR"
repeat_region 4774. .7112
/rept.family="L1"
repeat_region 7123. .7652
/rept.family="L1"
repeat_region 7629. .7771
/rept.family="L1"
repeat_region 8144. .8445
/rept.family="Alu"
repeat_region 8569. .8595
/rept.family="AT_rich"
repeat_region 10485. .10544
/rept.family="MALR"
repeat_region 10566. .10845
/rept.family="MALR"
repeat_region 10846. .11208
/rept.family="MALR"
repeat_region 11212. .11955
/rept.family="L1"
repeat_region 11987. .12299
/rept.family="Alu"
repeat_region 12300. .12779
/rept.family="L1"
repeat_region 12789. .13247
/rept.family="L1"
repeat_region 13277. .13669
/rept.family="Alu"
repeat_region 13705. .13838
/rept.family="L1"
repeat_region 13822. .14095
/rept.family="MER2_type"
repeat_region 14095. .14190
/rept.family="MER2_type"
repeat_region 14191. .14346
/rept.family="L1"
repeat_region 16001. .16141
/rept.family="L2"
repeat_region 16207. .16245
/rept.family="MIR"
repeat_region 16734. .16773
/rept.family="(CAAAAA)n"
repeat_region 16922. .17761
/rept.family="L1"
repeat_region 17933. .17957
/rept.family="(CA)n"
repeat_region 17980. .18101
/rept.family="MIR"
repeat_region 18102. .18465
/rept.family="MALR"
repeat_region 18466. .18516
/rept.family="MIR"
repeat_region 19440. .19313
/rept.family="CT-rich"
repeat_region 21201. .21295
/rept.family="L2"
repeat_region 21297. .21835
/rept.family="ERVL"
repeat_region 27445. .27997
/rept.family="ERVL"
repeat_region 28051. .28091
/rept.family="AT_rich"
repeat_region 28096. .28202
/rept.family="Alu"
repeat_region 28210. .28246
/rept.family="(TAAA)n"
repeat_region 28295. .28319
/rept.family="(TG)n"
repeat_region 28324. .28779
/rept.family="L1"
repeat_region 28780. .29214
/rept.family="MALR"
repeat_region 29215. .29704
/rept.family="L1"
repeat_region 29707. .30227
/rept.family="L1"
repeat_region 30217. .30997
/rept.family="L1"
repeat_region 31381. .31798
/rept.family="L1"
repeat_region 31819. .31965
/rept.family="MER1_type"
repeat_region 33297. .33558
/rept.family="CRL"
repeat_region 34871. .34894
/rept.family="AT_rich"
repeat_region 41290. .41408
/rept.family="MIR"
repeat_region 41745. .42317
/rept.family="ERVL"
repeat_region 42548. .42714
/rept.family="MIR"
repeat_region 42715. .42737
/rept.family="AT_rich"
repeat_region 44278. .44461
/rept.family="MIR"
repeat_region 44508. .44542
/rept.family="AT_rich"
repeat_region 45741. .45799
/rept.family="MALR"

Query Match 70.4%; Score 19; DB 9; Length 108100;
Best Local Similarity 81.5%; Pred. No. 3.1e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GCGGCCGAATTGAGTTATGATCTCT 27
Db 29004 GTCGAGAGTGAAGTTATGATCAT 29030

RESULT 23
CR407559 LOCUS CR407559 148587 bp DNA linear HTG 11-AUG-2004
DEFINITION Danio rerio clone CH211-283G2, WORKING DRAFT SEQUENCE.
ACCESSION CR407559
VERSION CR407559.7 GI:50978507
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Bukayot, Metzger, Chordata; Craniata; Vertebrata; Buteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE Lovell, J.
AUTHORS Direct Submission
TITLE Submitted (10-AUG-2004) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL Cambridge, CB10 1SA, UK. E-mail inquiries: Clonerquest@sanger.ac.uk
COMMENT zfish-help@sanger.ac.uk. Clone requests: Clonerquest@sanger.ac.uk
On Aug 4, 2004 this sequence version replaced gi:50724958.

```

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: zfish-help@sanger.ac.uk
 ----- Project Information
 Center project name: zc283G2
 ----- Summary Statistics
 Assembly program: XCAP4; version 4.5
 Chemistry: Dye-terminator; 100% of reads
 Consensus quality: 148583 bases at least Q40
 Consensus quality: 148587 bases at least Q30
 Consensus quality: 148587 bases at least Q20
 Insert size: 148587 sum-of-contigs
 Insert size: 163547; 10.1% error; agarose-fp
 Quality coverage: 11.56x in Q20 bases; sum-of-contigs Quality
 coverage: 11.33x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES source Location/Qualifiers

1..148587 /organism="Danio rerio"
 /mol type="genomic DNA"
 /db_xref="taxon:7955"
 /clones="CHA11-283G2"
 /clone lib="CHONI-211"
 1..148587 /note="assembly_fragment:01416"

ORIGIN

Query Match 70.4%; Score 19; DB 2; Length 148587;
 Best Local Similarity 81.5%; Pred. No. 3.2e+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GCGGCCGAATTGAAAGTTATGATCTCT 27
 Db 54365 GCGGCCGAATTGAAATCTATGACACCT 54391

RESULT 24 HSXKSRPXM HSXKSRPXM 175531 bp DNA linear PRI 14-DEC-1999

DEFINITION Homo sapiens chromosome X sequence from 5 cosmids, map Xp21.1

ACCESSION AL121577

VERSION AL121577.1 GI:5931934

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.

REFERENCE 1. Ramser,J., Borzym,K., Steffens,C., Klein,M., Monaco,A., Ho,M.,
 Reinhardt,R. and Lehrach,H.

JOURNAL Unpublished
 JOURNAL 2 (bases 1 to 175531)

REFERENCE 1. MPIMG.

AUTHORS Direct Submission

TITLE Submitted (27-SEP-1999) MPIMG, Abt. Lehrach, Max Planck Institut
 fuer Molekulare Genetik, Ihnestrasse 73, Berlin, 14195 Germany

COMMENT Contig 01 1..175531
 Clone ICRF104 3C29 received from the Resource Centre of the Human
 Genome Project at the Max-Planck-Institut for Molecular Genetics.
 Clones ICRF104 135D2, ICRF104 203D3, ICRF104 47E10, and ICRF104
 49A7 received from A. Monaco.

FEATURES source Location/Qualifiers

1..175531 /organism="Homo sapiens"
 /mol type="genomic DNA"
 /db_xref="taxon:9606"

source

1..442265 /organism="Homo sapiens"
 /mol type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="3C29"
 35604..372266 /organism="Homo sapiens"
 /mol type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="135D2"
 61035..103556 /organism="Homo sapiens"
 /mol type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="203D3"
 94955..134676 /organism="Homo sapiens"
 /mol type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="41B10"
 128388..175531 /organism="Homo sapiens"
 /mol type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="49A7"
 complement(11297..13484)
 repeat_region /rpt family="LINE/L1"
 2391..276 /note="LINE2"
 repeat_region /rpt family="LINE/L2"
 2824..3310 /note="MER68A"
 repeat_region /rpt family="Other/MER21_gro"
 3379..3484 /note="MLTIG"
 repeat_region /rpt family="LINE/L2"
 276 /note="LINE2"
 repeat_region /rpt family="LINE/L2"
 53336..53337 /note="MER5A"
 repeat_region /rpt family="DNA/MER1_type"
 6836..6964 /note="MLTIG"
 repeat_region /rpt family="LTR/MaLR"
 8576..8766 /note="MER"
 repeat_region /rpt family="SINE/MIR"
 8784..8810 /note="TAAATTA"
 repeat_region /rpt family="DNA/MER1_type"
 8813..9120 /note="AluSQ"
 9150..9258 /note="LINE2"
 repeat_region /rpt family="DNA/MER5A"
 9949..10244 /note="LINE2"
 repeat_region /rpt family="LINE/L1"
 11097..11273 /note="MER5B"

```

repeat_region          complement(36042. .36210)
                      /note="CATA)n"
                      /rpt_family="DNA/MER1_type"
                      complement(11818. .12052)
                      /note="MIR"
repeat_region          /rpt_family="SINE/MIR"
                      complement(12106. .12208)
                      /note="MER8A"
repeat_region          /rpt_family="DNA/MER1_type"
                      complement(112211. .12377)
                      /note="MIR"
repeat_region          /rpt_family="SINE/MIR"
                      complement(12211. .12088)
                      /note="AT rich"
repeat_region          /rpt_family="Low_complexity"
                      complement(13211. .13211
                      /note="AT rich"
repeat_region          /rpt_family="Low_complexity"
                      complement(13675. .13842
                      /note="MER8A"
repeat_region          /rpt_family="DNA/MER1_type"
                      complement(15296. .15743)
                      /note="MITC"
repeat_region          /rpt_family="LTR/MaLR"
                      complement(20444. .20557)
                      /note="(TAATA)n"
repeat_region          /rpt_family="Simple_repeat"
                      complement(20723. .20977)
                      /note="LIMD2"
repeat_region          /rpt_family="LINE/L1"
                      complement(21024. .21308)
                      /note="LIP5"
repeat_region          /rpt_family="LINE/L1"
                      complement(21434. .21464
                      /note="AT rich"
repeat_region          /rpt_family="Low_complexity"
                      complement(22607. .22671)
repeat_region          /rpt_family="SINE/MIR"
                      complement(23130. .23106
                      /note="TMIR"
repeat_region          /rpt_family="SINE/MIR"
                      complement(24453. .24442
                      /note="MBR5A"
repeat_region          /rpt_family="DNA/MER1_type"
                      complement(25305. .25308
                      /note="LIMC4"
repeat_region          /rpt_family="LINE/L1"
                      complement(25337. .31471)
                      /note="LIP22"
repeat_region          /rpt_family="LINE/L1"
                      complement(31474. .31823
                      /note="LIMC4"
repeat_region          /rpt_family="LINE/L1"
                      complement(32363. .32389
                      /note="AT rich"
repeat_region          /rpt_family="Low_complexity"
                      complement(32409. .32218)
                      /note="MIR"
repeat_region          /rpt_family="LINE/L1"
                      complement(32514. .32678
                      /note="LIMA5A"
repeat_region          /rpt_family="LINE/L1"
                      complement(32515)
                      /note="LIMA5A"
repeat_region          /rpt_family="LINE/L1"
                      complement(32409. .32218)
                      /note="LIP15"
repeat_region          /rpt_family="LINE/L1"
                      complement(34927. .35033
                      /note="GA)n"
repeat_region          /rpt_family="Simple_repeat"
                      complement(35511. .36012)
                      /note="LIME3"
repeat_region          /rpt_family="LINE/L1"

repeat_region          complement(36042. .36210)
                      /note="CATA)n"
                      /rpt_family="Simple_repeat"
                      complement(36700. .36731)
                      /note="CAAA)n"
repeat_region          /rpt_family="Simple_repeat"
                      complement(37034. .37386)
                      /note="TREC1"
                      /rpt_family="LTR/MaLR"
repeat_region          complement(37387. .37472
                      /note="LINE2"
                      /rpt_family="LINE/L2"
repeat_region          complement(38203. .38260)
                      /note="AT rich"
                      /rpt_family="Low_complexity"
repeat_region          complement(3842. .38176
                      /note="MIR"
repeat_region          complement(40343. .40406)
                      /note="MIR"
                      /rpt_family="SINE/MIR"
repeat_region          complement(40728. .40966)
                      /note="MIR"
repeat_region          complement(41055. .41237
                      /note="MIR"
                      /rpt_family="SINE/MIR"
repeat_region          Query Match    70.4%; score 19; DB 9;
Best Local Similarity 81.5%; pred. 3.3e+02; Matches 0; M:matches 5; Indels 0; Gaps 0;

Qy   1  GCGGCCAACGTGAAAGTTATCTCTT
Db  140719 GCTTCCTCCATGAGCTTGATCT 140745

RESULT 25
AC125238
LOCUS   Homo sapiens BAC Clone RP11-711C24
DEFINITION AC125238
ACCESSION AC125238.5
VERSION  GI:22267884
KEYWORDS HTG,
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
AUTHORS Nguyen, C., Bielicki, L. and Schatzkamer, K.
TITLE   The sequence of Homo sapiens BAC clone RP11-711C24
JOURNAL Unpublished (2001)
PUBMED  99063792
REFERENCE 2 (bases 1 to 196622)
AUTHORS Sulston, J. E. and Waterston, R.
TITLE   Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
PUBMED 9847074
REFERENCE 1 (bases 1 to 196622)
AUTHORS Waterston, R.H.
TITLE   Direct Submission
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
PUBMED 4 (bases 1 to 196622)
REFERENCE 2 (bases 1 to 196622)
AUTHORS Nguyen, C., Bielicki, L. and Schatzkamer, K.
TITLE   Direct Submission
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
PUBMED 5 (bases 1 to 196622)
REFERENCE 3 (bases 1 to 196622)
AUTHORS Waterston, R.
TITLE   Direct Submission
JOURNAL Submitted (08-AUG-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108 USA
PUBMED 4 (bases 1 to 196622)
REFERENCE 4 (bases 1 to 196622)
AUTHORS Waterston, R.H.
TITLE   Direct Submission
JOURNAL Submitted (08-AUG-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108 USA
PUBMED 5 (bases 1 to 196622)
REFERENCE 5 (bases 1 to 196622)
AUTHORS Waterston, R.
TITLE   Direct Submission
JOURNAL Submitted (16-AUG-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

```

REFERENCE

6 (bases 1 to 196622)

AUTHORS Waterston,R.
Direct Submission

JOURNAL Submitted (15-OCT-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

TITLE On Aug 16, 2002 this sequence version replaced gi:22138709.

COMMENT ----- Genome Center

Center: Washington University Genome Sequencing Center

Center Code: WUGSC

Web site: <http://genome.wustl.edu/gsc>

Contact: sapiens@wustl.wustl.edu

----- Summary Statistics

Center project name: H_NH0711C24

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Oboegwa,K., Woon,P.Y., Zhao,B., Frangen,E., Tateno,M., Cataneo,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome (BAC) libraries. Genomics, 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC069314 and AC010138.

A transposon was identified in the finished region of this clone and removed prior to submission. The transposon would insert after base pair 83916 of this sequence.

The sequence from base 156374 to base 156453 was derived from one plasmid subclone.

The sequence from base 156479 to base 156499 was derived from one plasmid subclone.

Polymorphisms have been identified between AC069314, AC010138 and this sequence.

Data from AC069314 and AC010138 was used to finish this clone.

FEATURES

Location/Qualifiers
1..19622
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="12"
/Clone="RPCI-711C24"
/clone.lib="RPCI-11"
repeat_region 215..236
/ rpt_family="TTTA" n"

repeat_region 240..520

/ rpt_family="Alu"

repeat_region 892..1066

/ rpt_family="L1"

repeat_region 1575..1885

/ rpt_family="Alu"

repeat_region 1991..2078

/ rpt_family="MIR"

repeat_region 2079..2387

/ rpt_family="Alu"

repeat_region 2418..2487

/ rpt_family="Alu"

repeat_region 2488..2764

/ rpt_family="Alu"

repeat_region 2771..2871

/ rpt_family="Alu"

repeat_region 3123..3190

/ rpt_family="L2"

repeat_region 3296..3399

/ rpt_family="Alu"

repeat_region 3612..3919

/ rpt_family="Alu"

repeat_region 3926..3988

/ rpt_family="MIR"

repeat_region 4192..4493

/ rpt_family="Alu"

repeat_region 5165..5256

/ rpt_family="AchHobo"

repeat_region 5257..5293

/ rpt_family="L1"

repeat_region 5599..5623

/ rpt_family="(TG)n"

repeat_region 5624..5748

/ rpt_family="Alu"

repeat_region 5752..5829

/ rpt_family="L1"

repeat_region 6043..6552

/ rpt_family="AchHobo"

repeat_region 6670..6803

/ rpt_family="L2"

repeat_region 6804..7097

/ rpt_family="Alu"

repeat_region 7098..7192

/ rpt_family="L2"

repeat_region 7193..7515

/ rpt_family="Alu"

repeat_region 7516..7745

/ rpt_family="L2"

repeat_region 7746..8054

/ rpt_family="Alu"

repeat_region 8055..8264

/ rpt_family="L2"

repeat_region 8265..8365

/ rpt_family="Alu"

repeat_region 8566..9242

/ rpt_family="L2"

repeat_region 9381..9582

/ rpt_family="Alu"

repeat_region 9688..9997

/ rpt_family="Alu"

repeat_region 10137..10384

/ rpt_family="L2"

repeat_region 10773..11075

/ rpt_family="Alu"

repeat_region 11076..11379

/ rpt_family="Alu"

repeat_region 11495..11580

/ rpt_family="MIR"

repeat_region 11927..12020

/ rpt_family="MIR"

repeat_region 12146..12237

/ rpt_family="MIR2_type"

repeat_region 12238..12534

/rpt family='Alu"						
repeat_region	12535..12682					
repeat_region	/rpt family="MER2_type"					
repeat_region	12663..13170					
repeat_region	/rpt family="MALR"					
repeat_region	13171..13193					
repeat_region	/rpt family="MER2_type"					
repeat_region	13631..14034					
repeat_region	/rpt family="MALR"					
repeat_region	14633..14966					
repeat_region	/rpt family="L1"					
repeat_region	14967..15250					
repeat_region	/rpt family="Alu"					
repeat_region	15251..15303					
repeat_region	/rpt family="L1"					
repeat_region	15796..16074					
repeat_region	/rpt family="Alu"					
repeat_region	16075..16315					
repeat_region						
Query Match	70 4%; Score 19; DB 9; Length 196622;					
Best Local Similarity	100.0%	Pred. No.	3.3e+02;			
Matches	19;	Conservative	0;	Mismatches	0;	
				Indels	0;	Gaps 0;
Y	9 AATTGAGTATGATGTATCCT 27					
b	119311 AATTGAGTATGATGTATCCT 119329					
RESULT	AC094790	DNA	linear	HTG 09-MAY-2003		
OCUS	235811 bp					
COMMENT	Rattus norvegicus clone CH230-4C5, WORKING DRAFT SEQUENCE, 2 unorderd pieces.					
ACCESSION	AC094790	GI:30466636				
VERSION	AC094790_7	HTGS_DRAFT; HTGS_FULLTOP.				
LEWDORS	HTGS_PHASEB; HTGS_NORVY rat					
SOURCE	Rattus norvegicus					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
1 (bases 1 to 235811)						
AUTHORS	Muzey,D,Marie.,Metzker,M.Lee.,Abrahazon,S.,Adams,C.,Alder,J.,Allen,C.,Allen,H.,Alisbrooks,S.,Amin,A.,Angulano,D.,Ayvagi,A.,Barber,M.,Barnstead,M.,Benahmed,F.,Baldwin,D.,Bandarainaki,D.,Barber,M.,Barnstead,M.,Benahmed,F.,Biswal,K.,Blair,J.,Blankenburg,K.,Bly,B.,Brown,M.,Bryant,N.,Buhay,C.,Burch,P.,Burrell,K.,Calderon,E.,Cariena,V.,Carter,K.,Cavazos,I.,Cesar,H.,Center,A.,Chacko,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,Cleveland,C.,Cockrell,R.,Cox,C.,Coylne,M.,Cree,A.,D'Souza,L.,David,M.L.,Davis,C.,Davy-Carroll,L.,De Andra,C.,Dederich,D.,Delgado,O.,Denson,S.,Deramo,C.,Ding,Y.,Dinh,H.,Divya,K.,Draper,H.,Dugan-Rocha,S.,Dunn,A.,Durbin,K.,Duyal,B.,Eavey,K.,Egan,A.,Escott,M.,Eugene,C.,Evans,C.A.,Fallas,T.,Fan,G.,Fernandez,S.,Finley,M.,Flagg,N.,Forbes,L.,Foster,M.,Fraser,C.M.,Gabis,A.,Ganta,R.,Garcia,A.,Garner,T.,Gazza,M.,Gebrgeorgis,E.,Gear,K.,Gill,R.,Grady,M.,Guerra,W.,Guevara,W.,Gumaratne,P.,Haaland,W.,Hamil,C.,Hamilton,C.,Hamilton,K.,Harvey,Y.,Harlak,P.,Hawes,A.,Henderson,N.,Hernandez,J.,Hernandez,R.,Hines,S.,Hladun,S.L.,Hodgson,A.,Hogues,M.,Hollins,B.,Howells,S.,Hulyk,S.,Hume,J.,Idlebird,D.,Jackson,A.,Jackson,L.,Jacob,L.,Jiang,H.,Johnson,B.,Johnson,R.,Jolivet,A.,Karpathy,S.,Kelly,S.,Khan,Z.,King,L.,Kovar,C.,Kowalewski,C.,Kraft,C.L.,Lebow,H.,Levan,J.,Lewis,L.,Li,Z.,Liu,J.,Lorenshuwa,L.,Liu,W.,Liu,Y.,London,P.,Longenecker,S.,Lopez,J.,Maheshwari,M.,Mahindarinne,M.,Mahnoud,M.,Malloy,K.,Mangum,A.,Mapua,P.,Martin,K.,Martinez,E.,Mawhinney,S.,McLeod,M.P.,McNeil,T.Z.,Meinen,E.,Milosavljevic,A.,Miner,G.,Minja,B.,Montemayor,J.,Moore,S.,Morris,M.,Morris,S.,Morris,S.,Munidasan,M.,Murphy,M.,Nair,L.,Nankervis,C.,Neal,D.,Newton,N.,Norris,S.,					
REFERENCE						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						
VERSION						
LEWDORS						
SOURCE						
ORGANISM						
1 (bases 1 to 235811)						
AUTHORS						
COMMENT						
ACCESSION						

clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:
The RPCI-13 BAC Library has been constructed by Kazutoyo Osegawa and Minako Tateno in the laboratory of Pieter de Jong (http://www.chori.org) from female C57Bl/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

source

```
/rpt_family="MaLR"
11700..12008
/rpt_family="RMR13A"
12588..12673
/rpt_family="Alu"
12714..13125
/rpt_family="L1"
13153..13269
/rpt_family="L1"
13786..13863
/rpt_family="L2"
13867..14082
/rpt_family="MER1_type"
14228..14307
/rpt_family="MaLR"
14385..14519
/rpt_family="Alu"
14525..14802
/rpt_family="MaLR"
15129..15431
/rpt_family="Alu"
15795..15941
/rpt_family="L1"
16555..16817
/rpt_family="L1"
17585..17932
/rpt_family="ERVK"
18027..18300
/rpt_family="B4"
18395..18732
/rpt_family="ERVK"
21585..22438
/rpt_family="ERVK"
22447..22774
/rpt_family="ERVK"
24311..24535
/rpt_family="MaLR"
24906..25152
/rpt_family="B4"
25951..26021
/rpt_family="B2"
27923..28031
/rpt_family="Alu"
28636..28756
/rpt_family="RMR15"
28777..28921
/rpt_family="B2"
28940..29227
/rpt_family="L1"
29257..29388
/rpt_family="L1"
28777..28905
/rpt_family="B2"
35031..35378
/rpt_family="L1"
36446..36534
/rpt_family="Alu"
36535..36918
/rpt_family="ERVK"
36919..36980
/rpt_family="ERVK"
38916..39136
/rpt_family="Alu"
37306..37382
/rpt_family="ERVK"
37682..38305
/rpt_family="ERVK"
38438..38670
/rpt_family="ERVK"
38916..39136
/rpt_family="ERVK"
39750..40336
/rpt_family="L1"
40249..40443
/rpt_family="ERVK"
repeat_region
11569..11600
```


ORIGIN

Query Match Similarity 68.9%; Score 18.6; DB 2; Length 218328;
Best Local Matches 84.0%; Pred. No. 5.4e+02; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GGCGCANTGAGTTATGATCTTCT 27
Db 9436 GGTCACANTGAGTTATGATCTTCT 9412

RESULT 3.0
AC141426 AC141426 227835 bp DNA linear ROD 15-MAY-2004
DEFINITION *Mus musculus* BAC clone RP23-280I14 from chromosome 12, complete sequence.

ACCESSION AC141426
VERSION AC141426.3
KEYWORDS FGT.
ORGANISM *Mus musculus* (house mouse)

REFERENCE 1 (bases 1 to 227835)
AUTHORS Tranil J., Bielicki L., and Haakenson, W.
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 227835)
AUTHORS McPherson, J.D. and Waterston, R.H.
JOURNAL Direct Submission (14-MAR-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 3 (bases 1 to 227835)
AUTHORS McPherson, J.D. and Waterston, R.H.
JOURNAL Direct Submission (15-MAR-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 227835)
AUTHORS Wilson, R.K.
JOURNAL Direct Submission (12-MAY-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 227835)
AUTHORS Wilson, R.K.
JOURNAL Direct Submission (15-MAY-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT On May 12, 2004 this sequence version replaced gi_28973968.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@watson.wustl.edu
----- Summary Statistics CG
Center project name: M_BA0280I14

<http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Osegawa and Minako Tateno in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>.

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC110175 and AC124414.

FEATURES	source	Location/Qualifiers
repeat_region	1..227835	/clone lib="RPCI-23" /rpt. family="B2"
repeat_region	1075..1150	/rpt. family="genomic DNA" /db_xref="taxon:10090" /chromosome="12" /map="12" /clone="RP23-280I14"
repeat_region	2008..2369	/rpt. family="CR1"
repeat_region	3807..3998	/rpt. family="B4"
repeat_region	6431..6575	/rpt. family="Alu"
repeat_region	6815..6945	/rpt. family="Alu"
repeat_region	8950..9174	/rpt. family="L1"
repeat_region	10003..10201	/rpt. family="B2"
repeat_region	11318..11388	/rpt. family="MER1-type"
repeat_region	12349..12551	/rpt. family="MTR"
repeat_region	13890..14038	/rpt. family="Alu"
repeat_region	21006..21163	/rpt. family="BVK"
repeat_region	21629..22113	/rpt. family="M4R"
repeat_region	22256..22359	/rpt. family="B4"
repeat_region	22451..22612	/rpt. family="MER1-type"
repeat_region	24695..24794	/rpt. family="L1"
repeat_region	25406..25514	/rpt. family="B4"
repeat_region	31802..31877	/rpt. family="B4"
repeat_region	33182..33267	/rpt. family="Alu"
repeat_region	33752..33805	/rpt. family="Alu"
repeat_region	33805..33885	/rpt. family="ERV1"
repeat_region	35298..35505	/rpt. family="B4"
repeat_region	36176..36447	/rpt. family="B2"
repeat_region	36371..36608	/rpt. family="B2"
repeat_region	36673..36863	

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Web Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see

```

repeat_region   / rpt_family="Alu"          68.9%; Score 18.6; DB 10; Length 227835;
repeat_region   36862 .36529              Best Local Similarity 84.0%; Pred. No. 5.e+02; Indels 0; Gaps 0;
repeat_region   / rpt_family="ID"          Matches 21; Conservative 0; Mismatches 4;
repeat_region   37030 .3724               repeat_region   63501 .64000
repeat_region   / rpt_family="L1"          repeat_region   63041 .63095
repeat_region   41720 .42055              repeat_region   64627 .65027
repeat_region   / rpt_family="MaLR"        repeat_region   63041 .63095
repeat_region   42603 .42256              repeat_region   65192 .65317
repeat_region   / rpt_family="B4"          repeat_region   63105 .63157
repeat_region   43610 .43702              repeat_region   63358 .65557
repeat_region   / rpt_family="L1"          repeat_region   67854 .68009
repeat_region   44106 .44176              repeat_region   / rpt_family="MER1_type"
repeat_region   44800 .45073              repeat_region   / rpt_family="B4"
repeat_region   45278 .45156              repeat_region   45278 .45156
repeat_region   / rpt_family="MaLR"        repeat_region   48693 .48863
repeat_region   49778 .50349              repeat_region   51582 .51669
repeat_region   / rpt_family="L1"          repeat_region   53188 .53186
repeat_region   50668 .51581              repeat_region   56043 .56098
repeat_region   / rpt_family="ERV1"        repeat_region   56325 .56403
repeat_region   51582 .51669              repeat_region   57963 .58029
repeat_region   / rpt_family="Alu"          repeat_region   58031 .58151
repeat_region   59118 .59288              repeat_region   59466 .59732
repeat_region   / rpt_family="L1"          repeat_region   60788 .61066
repeat_region   61086 .61009              repeat_region   61086 .61009
repeat_region   / rpt_family="B4"          repeat_region   61962 .62111
repeat_region   62202 .62475              repeat_region   62202 .62475
repeat_region   / rpt_family="B4"          repeat_region   62769 .62826
repeat_region   63105 .63157              repeat_region   62835 .62931
repeat_region   / rpt_family="MER1_type"  repeat_region   63041 .63095
repeat_region   / rpt_family="L1"          repeat_region   63041 .63095
repeat_region   65192 .65317              repeat_region   65192 .65317
repeat_region   / rpt_family="MER1_type"  repeat_region   65358 .65557
repeat_region   67854 .68009              repeat_region   / rpt_family="B4"

```

Query Match 68.9%; Score 18.6; DB 10; Length 227835;
 Best Local Similarity 84.0%; Pred. No. 5.e+02; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 4;

Db 40133 GCGAGCTTATGAGTTATCAATCCT 40157

RESULT 31 AC094794/C

AC094794 LOCUS Rattus norvegicus clone CH230-4N15, *** SEQUENCING IN PROGRESS ***.

DEFINITION HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.

VERSION AC094794 GI:30466624

KEYWORDS Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 227929)

AUTHORS Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J., Allen, C., Alien, H., Alsbrooks, S., Amin, A., Anguloano, D., Aryal, B., Ayagi, A., Ayodeji, I. M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstied, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, R., Cox, C., Coyne, M., Cree, A., D'Souza, L., Davila, M.L., Davis C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Derano, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Eaves, K., Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, T., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gabregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, B., Hulyk, S., Hume, J., Idlebird, D., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowalski, C., Kraft, C.L., Lebow, H., Lewis, J., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensubawa, L., Loulseged, H., Lorado, R.J., Lu, X., Ma, J., Manshewari, M., Mahindarane, M., Mahmoud, M., Mallory, K., Mangum, A., Mansum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mauhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, B., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidaswa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakelameh, O., Okwuonu, G., Olarnpunpongsoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, P., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puerto, M., Quiroz, J., Rachin, B., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Saverty, G., Scheer, G., Scott, G., Shatsman, S., Shen, H., Shetty, J., Saveriyeen, A., Sisson, I., Sitter, C.D., Smalls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Williamson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 227929)
 AUTHORS Worley, K.C.
 TITLE Direct Submission
 JOURNAL Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One

REFERENCE	3 (bases 1 to 227929)	Baylor Plaza, Houston, TX 77030, USA
AUTHORS	Rat Genome Sequencing Consortium.	
TITLE	Direct Submission	
JOURNAL	Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	
COMMENT	The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.	
FEATURES		
source	/organism="Rattus norvegicus"	
	/mol_type="genomic DNA"	
	/db_xref="Taxon:10116"	
	/clone="CH230-4N15"	
misc_feature	1..3288	
	/note="wgS end-extension	
	clone_end:596"	
	595..6182	
	/note="clone boundary	
	clone_end:spf	
	site:EcoRI	
	end_sequence:BH309552"	
ORIGIN		
	Query Match Similarity 68.9%	Score 18.6; DB 2;
	Best Local Matches 84.0% 21; Conservation	Length 227929;
Oy	0; Mismatches 0;	Indels 0; Gaps 0;
Db	3 GGCGCAATTGAACTTATCTATCCCT 27	3 GGCTACAAATGAAGTTTGTCCT 131857

TITLE	Direct Submission	misc_feature	/note="wgs contig"
JOURNAL	Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	misc_feature	120251_121880 /note="wgs contig"
COMMENT	The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.	misc_feature	163473_164837 /note="wgs contig"
FEATURES		ORIGIN	
source		Query Match	68.9%; Score 18.6; DB 2;
		Best Local Similarity	Pred. No. 5.4e+02;
		Matches	4; Mismatches 0;
		21; Conservative	Indels 0;
			Gaps 0;
		Qy	3 GCGCCGATTGAGTGTATGATCCT 27
		Db	86239 GCGCGTTGTAGTGTATGATCCT 86263
		RESULT 33	
		AC100967	247100 bp DNA linear HTG 11-AUG-2004
		LOCUS	Mus musculus chromosome 5 clone RP23-74L20 map 5, *** SEQUENCING IN PROGRESS ***, 9 unordered pieces.
		DEFINITION	
		AC100967	AC100967_10 GI:51101176
		ACCESSION	HTG; HTGS PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
		VERSION	
		KEYWORDS	
		SOURCE	Mus musculus (house mouse)
		ORGANISM	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
		REFERENCE	1 (bases 1 to 247100)
		AUTHORS	Birren,B., Nusbaum,C. and Lander,E.
		TITLE	Mus musculus chromosome 5, clone RP23-74L20
		JOURNAL	Unpublished
		REFERENCE	2 (bases 1 to 247100)
		AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Al.I.A., Allen,N., Boukhgalaite,B., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalaite,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazarro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Debellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,W., Graham,L., Grand-Pierre,N., Hagos,B., Heatord,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lebockzy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Minova,T., Mitra,V., Murphy,T., Narlitor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Repta,R., Riley,R., Rogov,P., Roy,A., Santos,R., Schauer,S., Schupback,R., Roman,J., Rosette,M., Seaman,S., Seaver,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straub,N., Subramanian,J., Talamas,J., Testa,J., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Vieil,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
		TITLE	Direct Submission
		JOURNAL	Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
		REFERENCE	3 (bases 1 to 247100)
		AUTHORS	Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Archichi,H.M., Barni,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalaite,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corrum,B., DeBellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Grand-Pierre,N., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Huie,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Menau,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicoli,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,

Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
 Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N.,
 Stojanovic, N., Stubbs, M., Talianas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Vel, R.,
 Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembret, L.,
 Zimmer, A. and Zody, M.
TITLE
 Direct Submission
JOURNAL
 Submitted (11-AUG-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT
 On Aug 11, 2004 this sequence version replaced g1:51036443.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
DEFINITION
 Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@broad.mit.edu
Project Information
Center project name: L14656
Center clone name: 74_L_20
FEATURES
source
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 9 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 16605: contig of 16605 bp in length
 * 167006 16705: gap of unknown length
 * 36350: contig of 19645 bp in length
 * 36351: 36350: gap of unknown length
 * 36351: 6038: contig of 23589 bp in length
 * 6038: gap of unknown length
 * 60339 60338: contig of 5791 bp in length
 * 60139 65329: gap of unknown length
 * 65329 66329: gap of unknown length
 * 66330 95301: contig of 29072 bp in length
 * 95301: 95301: gap of unknown length
 * 95302 122730: contig of 32229 bp in length
 * 122730 127510: gap of unknown length
 * 127431 133477: contig of 5947 bp in length
 * 127531 133478: gap of unknown length
 * 133478 133477: gap of unknown length
 * 133478 168446: contig of 34869 bp in length
 * 168446 168446: gap of unknown length
 * 168447 247100: contig of 78554 bp in length.
ORIGIN
source
 Query Match 68.9%; Score 18.6; DB 2; Length 247100;
 Best Local Similarity 84.0%; Pred. No. 5.4e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
DEFINITION
 Ratmus norvegicus clone CH210-62C13, *** SEQUENCING IN PROGRESS
RESULT
 AC129421 290066 bp DNA linear HTG 20-NOV-2002
LOCUS
 Ratmus norvegicus clone CH210-62C13, 5 unpaired pieces
DEFINITION
 AC129421 290066 bp DNA linear HTG 20-NOV-2002

Of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 146945)

REFERENCE

Worley, K.C.

Authors

Title

Journal

Submitted (31-JAN-2002)

Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 146945)

Worley, K.C.

Direct Submission

Submitted (31-JAN-2003)

Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Jan 31, 2003 this sequence version replaced gi:6492471.

Information: <http://www.hgsc.bcm.tmc.edu/> or email

sc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of Sequence Similarity are identified by BLAST (Nuc. Acids Res. 25:1389-1402) similarity (expect <1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

Location/Qualifiers

1. .146945
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="ttxon:9606"

FEATURES

source

/clone=TRP11-19E18"
757 .883
/standard_name="WI-16568"
811 .966
/standard_name="A010A40"

repeat_region

2810 .2849
/rpt_family="(A)n"
/chromosome="12"

repeat_region

3475 .3512
/rpt_family="T-rich"
3507 .3632

repeat_region

/standard_name="WI-13179"
3511 .3648
/standard_name="WI-145421"
complement("059 .4362")
/rpt_family="AluSc"

```

repeat_region          4890 .4922 /rpt_family="AT_rich"
repeat_region          5231 .5853 /rpt_family="LIME3A"
repeat_region          5963 .6217 /rpt_family="AluSg"
repeat_region          6797 .7094 /rpt_family="AluJo"
repeat_region          7304 .7614 /rpt_family="AluY"
repeat_region          7630 .8065 /rpt_family="HAL1"
repeat_region          8132 .8366 /rpt_family="HAL1"
repeat_region          8413 .8724 /rpt_family="HAL1"
repeat_region          8868 .8941 /rpt_family="AluX"
repeat_region          9016 .9163 /rpt_family="(TA)n"
repeat_region          9785 .9805 /rpt_family="AluSc"
repeat_region          9785 .9805 /rpt_family="AT_rich"
repeat_region          9813 .9848 /rpt_family="LIMB8"
repeat_region          9849 .10139 /rpt_family="AluSx"
repeat_region          1040 .10277 /rpt_family="LIMB8"
repeat_region          10304 .10680 /rpt_family="LIMB3"
repeat_region          10681 .10927 /rpt_family="LIMB3"
repeat_region          10933 .10984 /rpt_family="LIME1"
repeat_region          12687 .12937 /rpt_family="L1"
repeat_region          12941 .13226 /rpt_family="AluSg"
repeat_region          13205 .13393 /rpt_family="L2"
repeat_region          13526 .13990 /rpt_family="L2"
repeat_region          14980 .15014 /rpt_family="AT_rich"
repeat_region          15365 .15385 /rpt_family="(GA)n"

Query Match          68.1% Score 18.4; DB 9;
Best Local Similarity 95.0%; Pred. No. 6.6e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy                  8 CAATTGAGTTATGATCTCT 27
Db                  43751 CATTGGGTTATGATCTCT 43732

RESULT 36          AC110187/c
LOCUS          AC110187
DEFINITION      Mus musculus chromosome 7 clone RP24-541E23 map 7, WORKING DRAFT
ACCESSION      AC110187
VERSION       AC110187.5 GI:45430174

```

KEYWORDS	HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE	Mus musculus (house mouse)
ORGANISM	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus .
REFERENCE	1 (bases 1 to 18054)
AUTHORS	Birren, B., Nusbaum, C. and Lander, E.
TITLE	Unpublished
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 18054)
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,' Brown, A., Barna, N., Bastin, V., Boguslavsky, I.L., Boukhgalter, B.,' Brown, A., Camarata, J.J., Campopiano, A., Chang, J.J., Chazaro, B., Choepel, I.Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooks, P., DeAngelis, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardya, S., Ginde, S., Gord, S., Goyetz, M., Graham, L., Grand-Pierre, N., Hagos, S., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Keilis, C., Larocque, K., Lanazares, R.,' Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,' Landers, T., Lehozcy, J., Levine, R., Liu, G., Maclean, C.,' Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,' McEwan, P., McKernan, K., Meldrim, J., Menas, J.B., Mihova, T.,' Mlenga, V., Murphy, T., Navlak, J., Nguyen, C., Nicol, R., Norbu, C.,' Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, J., Oliver, J.,' Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,' Retta, R., Rieback, M., Riley, R., Rose, C., Rogov, P., Roman, J.,' Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,' Seaman, S.,' Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,' Strauss, N., Subramanian, A., Talama, J., Tesfaye, S., Theodore, J.,' Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,' Vie, L.R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,' Zainoun, J., Zembek, L., Zimmer, A., and Zody, M.,'
TITLE	Direct Submission
JOURNAL	Submitted (10-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	3 (bases 1 to 18054)
AUTHORS	Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,' Anderson, M., Arachchi, H.M., Barna, N., Bastin, V., Bloom, T.,' Boguslavsky, I.L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,' Collymore, A., Cook, A., Cooke, P., Corum, B., DeAngelis, K.,' Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,' Ferreira, P., Fitzgerald, M., Gage, D., Galaga, J., Gardya, S.,' Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,' Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, J., Jones, C.,' Kamat, A., Karatas, A., Keilis, C., Landers, T., Levine, R.,' Lindblad-Toh, K., Liu, X., Liu, A., Mabritt, R., Maclean, C.,' Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,' Meldrim, J., Menas, J.B., Mihova, T.,' Mengen, V., Murphy, T.,' Naylor, J.,' Nguyen, C.,' Nicol, R.,' Norbu, C.,' O'Connor, T.,' O'Donnell, P.,' O'Neill, D.,' Oliver, J.,' Peterson, K.,' Phunkhang, P.,' Pierre, N.,' Rachapka, A.,' Ramasamy, U.,' Raymond, C.,' Retta, R.,' Rie, C.,' Roman, J.,' Schauer, S.,' Schupback, R.,' Seaman, S.,' Severy, P.,' Smith, C.,' Spencer, B.,' Stange-Thomann, N.,' Stojanovic, N.,' Stubbs, M.,' Talama, J.,' Tesfaye, S.,' Theodore, J.,' Topham, K.,' Travers, M.,' Vassiliev, H.,' Venkataraman, V.S.,' Vie, L.R.,' Vo, A.,' Wilson, B.,' Wu, X.,' Wyman, D.,' Young, G.,' Zainoun, J.,' Zembek, L.,' Zimmer, A., and Zody, M.,'
TITLE	Direct Submission
JOURNAL	Submitted (13-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Mar 13, 2004 this sequence version replaced gi:20503122. All repeats were identified using RepeatMaster: Smit, A.A.F. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMaster.html
-----	Genome Center
-----	Center: Whitehead Institute/ MIT Center for Genome Research Center Code: WIBR Web site: http://www-seq.wi.mit.edu
-----	Contact: sequence_submissions@genome.wi.mit.edu
-----	Project Information
-----	Center project name: 119479
-----	Center clone name: 541_E_23
-----	Summary Statistics Sequencing vector: Plasmid; n/a; 100% of reads


```

misc_feature
  /rpt_family="Lx"
  3638 - 3643
  /note="clone boundary
clone_end:=P6
site_ECORI"
repeat_region
  6732 - 6781
  /rpt_family="AT_rich"
  complement(8778 - 10152)
/rpt_family="ORRI1A-int"
repeat_region
  10153 - 10175
  /rpt_family="(CAAAA)n"
complement(10176 - 10400)
repeat_region
  11476 - 11506
  /rpt_family="ORRI1A-int"
complement(10401 - 10736)
/rpt_family="ORRI1A2"
repeat_region
  11887 - 11910
  /rpt_family="AT_rich"
complement(113543 - 13815)
repeat_region
  114227 - 11584
  /rpt_family="AT_rich"
/rpt_family="L1"
repeat_region
  14427 - 14448
  /rpt_family="AT_rich"
/rpt_family="L1"
repeat_region
  15058 - 15573
  /rpt_family="AT_rich"
/rpt_family="Lx5"
repeat_region
  15575 - 15754
  /rpt_family="ORRI1A2"
/rpt_family="L1"
repeat_region
  15755 - 15773
  /rpt_family="L1"
repeat_region
  17980 - 18892
  /rpt_family="L1VL2"
repeat_region
  18891 - 20598
  /rpt_family="L1_RN"
/rpt_family="L1"
repeat_region
  20686 - 21519
  /rpt_family="L1"
repeat_region
  21951 - 23084
  /rpt_family="L1"
repeat_region
  23085 - 23122
  /rpt_family="(CA)n"
/rpt_family="Lx9"
repeat_region
  23123 - 24060
  /rpt_family="L1"
repeat_region
  24062 - 28131
  /rpt_family="L1"
repeat_region
  28132 - 28354
  /rpt_family="ETnERBV2"
/rpt_family="L1_MM"
repeat_region
  28734 - 28934
  /rpt_family="AT_rich"
complement(28975 - 29087)
repeat_region
  29088 - 29471
  /rpt_family="Lx9"
/rpt_family="B3"
repeat_region
  29563 - 29749
  /rpt_family="Lx4"
repeat_region
  29750 - 29773
  /rpt_family="AT_rich"
complement(30096 - .30187)
/rpt_family="B3"
repeat_region
  30187 - 30400
  /rpt_family="B3"
complement(30197 - .30400)
/rpt_family="Lx5"
repeat_region
  32153 - 34475
  /rpt_family="Lx2"
repeat_region
  34476 - 34500
  /rpt_family="AT_rich"
complement(34501 - .34500)
/rpt_family="B2_Mm2"
repeat_region
  34659 - 34808
  /rpt_family="Lx2"
repeat_region
  36658 - 36658
  /rpt_family="Lx2"
repeat_region
  36654 - 37726
  /rpt_family="Lx2"
complement(37781 - .37846)
repeat_region
  37855 - 37907
  /rpt_family="L1_MM"
complement(39361 - .39633)
/rpt_family="MTD"
repeat_region
  39771 - 40347
  /rpt_family="MERVL"
repeat_region
  40346 - 40666
  /rpt_family="MERVL"
repeat_region
  40691 - 41005
  /rpt_family="Lx5"
complement(41009 - .41358)

repeat_region
  complement(34807 - .35361)
/rpt_family="Lx2"
complement(35596 - .36658)
repeat_region
  36654 - 37726
  /rpt_family="Lx2"
repeat_region
  37781 - .37846
repeat_region
  37855 - 37907
  /rpt_family="(CAAAA)n"
repeat_region
  37927 - .39279
  /rpt_family="Lx2"
repeat_region
  39361 - .39633
  /rpt_family="MTD"
repeat_region
  39771 - 40347
  /rpt_family="MERVL"
repeat_region
  40346 - 40666
  /rpt_family="MERVL"
repeat_region
  40691 - 41005
  /rpt_family="Lx5"
complement(41009 - .41358)

repeat_region
  complement(34807 - .35361)
/rpt_family="Lx2"
complement(35596 - .36658)
repeat_region
  36654 - 37726
  /rpt_family="Lx2"
repeat_region
  37781 - .37846
repeat_region
  37855 - 37907
  /rpt_family="GAAAA)n"
repeat_region
  37927 - .39279
  /rpt_family="Lx2"
repeat_region
  39361 - .39633
  /rpt_family="MTD"
repeat_region
  39771 - 40347
  /rpt_family="MERVL"
repeat_region
  40346 - 40666
  /rpt_family="MERVL"
repeat_region
  40691 - 41005
  /rpt_family="Lx5"
complement(41009 - .41358)

Query Match          68.1%
Best Local Similarity 95.0%
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      8 CAATTGAGTATGATGATCCT 27
Db      89552 CAATTGAACTATGATGATCCT 89571

RESULT 38
AC107233/C
LOCUS   199432 bp    DNA  LINEAR HTG 04-MAR-2003
DEFINITION Mus musculus clone RP23-151c4, WORKING DRAFT SEQUENCE, 11 unordered
PIECES.
AC107233
VERSION AC107233.4 GI:28827989
HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS Mus musculus (house mouse)
SOURCE
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 199432)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-151c4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 199432)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Boukhgalaer,B.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Chang,J., Chazaro,B.,
Brown,A., Camarata,J., Campopiano,A., Collingore,A., Cook,A., Chopek,K., Colangelo,M., Collins,S., Dodge,S., Faro,S., Cooke,P., Dearlano,K., Devar,K., Diaz,J.S., Galagan,J., Gardyna,S., Ferreira,P., FitzHugh,W., Gage,D., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heatord,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lazarares,R., Landers,B., Leibovitz,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N., McPheeers,R., Meldrin,J., McCarthy,M., McEwan,P., McKernan,K., Murphy,T., Nayor,J., Nguyen,C., Menes,L., Mihova,T., Mlenga,V., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Petersen,K., Phunkhang,P., Pierre,N., Pollara,V., Raymon,C., Recart,R., Rieback,M., Riley,R., Riss,C., Rogov,P., Schupback,R., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,S., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission
JOURNAL Submitted (17-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

```


Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabis,A., Gao,J., Garcia,A., Garner,T., Gatzza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,X., Hart,M., Havlik,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Hollaway,C., Hollins,B., Homai,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.B., Karlsson,B., Jia,Y., Johnson,R., Jolivet,S., Jouah,S., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louiseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucifer,R., Luna,R., Ma,J., Maheshwari,M., Mappa,P., Martin,R., Martindale,A., Martinez,E., Massy,E., Mawhinney,B., McLeod,M., Meador,M., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moyer,M., Neal,D., Newton,J., Nguyen,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokonwo,S., Oghu,M., Okwuonu,G., Oragny,N., Oriedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,B., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubano,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshtari,N., Sisson,I., Sodergren,B., Sonaikar,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamersa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Umarani,K., Vasquez,I., Ver,V., Villalon,D., Vinson,R., Wang,Q., Williams,G., Williamson,A., Wleczek,R., Wooden,S., Watlington,K., Wu,C., Wu,Y., Wu,Y.P., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G., and Gibbs,B.R.

TITLE Direct Submission
JOURNAL Unpublished
AUTHORS Worley,K.C.

TITLE Direct Submission
JOURNAL Submitted (15-SEP-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Apr 24, 2002 this sequence version replaced gi:14547712.
REFERENCE Authors Worley,K.C.

TITLE Direct Submission
JOURNAL JOURNAL OF MOLECULAR AND HUMAN GENETICS, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Drafting Center Code: BCM
 Contact: hgsc-help@bcm.tmc.edu
REFERENCE Authors Worley,K.C.

TITLE Direct Submission
JOURNAL JOURNAL OF MOLECULAR AND HUMAN GENETICS, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT Center: Baylor College of Medicine
 Center clone name: RP1-551G13
 Sequencing vector: Plasmid;
 Sequencing vector: M13;
 Chemistry: Dye-primer Bodipy; 46% of reads
 Assembly program: Phrap; version 0.9029
 Consensus quality: 234054 bases at least Q40
 Consensus quality: 236536 bases at least Q30
 Consensus quality: 238152 bases at least Q20
 Estimated insert size: 250093; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
 * NOTE: This sequence may represent more than one clone.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 9 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence

* as soon as it is available and the accession number will
 * be preserved.
 * * 3042: contig of 3042 bp in length
 * * 3043 3141: gap of unknown length
 * * 3143 6130: contig of 2988 bp in length
 * * 6131 6230: gap of unknown length
 * * 6231 14402: contig of 8172 bp in length
 * * 14403 14502: gap of unknown length
 * * 14503 31187: contig of 16685 bp in length
 * * 31188 31287: gap of unknown length
 * * 31288 55865: contig of 24578 bp in length
 * * 55866 55965: gap of unknown length
 * * 55966 78599: contig of 23634 bp in length
 * * 78600 78699: gap of unknown length
 * * 78700 108718: contig of 30019 bp in length
 * * 108719 108818: gap of unknown length
 * * 108819 159929: contig of 51111 bp in length
 * * 159930 160029: gap of unknown length
 * * 160030 242301: contig of 82272 bp in length.
FEATURES
 source
 1. .242301
 1. .242301;
 /organism="Homo sapiens"
 /mol type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="12"
 /clone="RP11-551G13"
ORIGIN
 Query Match 68.1%; Score 18.4.; DB 2;
 Best Local Similarity 95.0%; Pred. No. 6.9e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 8 CAATTGAGTATGTTACCT 27
 Db 58390 CATTGAGGTATGTTACCT 58371
RESULT 40
SOLEU2
LOCUS SOLEU2
DEFINITION S.occidentalis LEU2 gene.
ACCESSION X79823
VERSION X79823.1 GI:791136
KEYWORDS 3-isopropylmalate dehydrogenase; LEU2 gene.
SOURCE Debaromyces occidentalis
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomyces; Saccharomycetales; Saccharomycetaceae; Debaromyces.
REFERENCE Iserentant,D. and Verachtert,H.
AUTHORS Iserentant,D.M.M.
TITLE Cloning and sequencing of the LEU2 homologue gene of Schwanomyces
 Yeast 11 (5), 467-473 (1995)
JOURNAL MEDLINE 95321019
PUBMED 7597851
FEATURES
 source
 1. .1465
 /organism="Debaromyces occidentalis"
 /mol type="genomic DNA"
 /strain="ATCC26077/R91"
 /db_xref="taxon:27300"
 gene
 144 . 1283
 /gene="LEU2"
 144 . 1283
 /gene="LEU2"
 /EC_number="1.1.1.85"
 /codon_start=1
 CDS

/product="3-isopropylmalate dehydrogenase"
/protein_id="CAA56224_1"
/db_xref="IGI:791137"
/db_xref="Swiss-Prot: P48012"
/translation="MELISYTAAPSSTYEPDQIRNQLDITLNEILNQVNTNYMGLVN
LLIQNFETPLRQVLSTGIGSDMTRINQVFPTIDEITRINCILHDLSNKVSRYCVEI
IRVGGMILPYQKKAFTRHEANLKLMTTRUNKFSKNSKTSKFLSNNLINKENYSTERIDS
IVSGSLAEELPKRLILKRLHALESERKLKGGKFENFDNETSILTYVNTLEEVIDAF
GCCOGSEBDRKIDSOURLVFTPTGKILTELATKWPPLQYGWLSRKVGVYELRNVPKGS
FFDLELIIIESDHLLPLETIEBSYYIRKKNEPIKSLSVDLMSILSLYNEKPLPVLSLL
PSMEVNSWCNLNEVITSYLGVSHTSSKQDFLRFLNVSKSGFKS"

ORIGIN

Query	Match	Score 67.4%	Score 18.2;	DB 8;	Length 1465;
Best	Local	Similarity 87.0%;	Prod. No. 5.5e+02;	Indels 0;	Gaps 0;
Matches 20;	Conservative 0;	Mismatches 3;			

Qy 5 CCGCAATTGAAAGTTATGTATCCCT 27
Db 772 CCGCAATTGAAAGTTATGTATGCT 794

Search completed: January 7, 2005, 11:33:30
Job time : 938.789 secs

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: January 7, 2005 10:58:39 ; Search time 205.579 Seconds (without alignments)

Title: US-09-786-502A-7

Perfect score: 27

Sequence: 1 gccccgcaattggaaatgtatcc 27

Scoring table: IDENTITY_NUC Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 82697722

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

Database :	N_Geneseq_23Sep04:*
1:	Geneseqn1980s:*
2:	Geneseqn1990s:*
3:	Geneseqn2000s:*
4:	Geneseqn2001ab:*
5:	Geneseqn2001bs:*
6:	Geneseqn2002ab:*
7:	Geneseqn2002bs:*
8:	Geneseqn2003ab:*
9:	Geneseqn2003bs:*
10:	Geneseqn2004cs:*
11:	Geneseqn2004ds:*
12:	Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description	
1	27	100.0	27	3	AAA10270		Aaa10270 Human CD2	
2	27	100.0	27	6	ABA92024		Abes9024 Human CD2	
3	19	70.4	328	12	ADL67231		Adl67231 Human CD2	
4	19	70.4	660	2	AAAT96357	CD28	CdNA	
5	19	70.4	663	9	ADA02992	Human CD2	Adada02992 Human CD2	
6	7	19	70.4	663	10	ADB72730		Adb72730 Human CD2
7	19	70.4	663	10	ADC55472		Adc85472 Human Cd2	
8	19	70.4	663	12	ADM74587		Adm74587 Human car	
9	19	70.4	762	2	AAQ28837		Aaq28837 Sequence	
10	19	70.4	1514	2	AAQ21167		Aaq21167 Human CD2	
11	19	70.4	1514	2	AAT36257		Aat36257 Human CD2	
12	19	70.4	1514	2	AAT14107		Aat14107 Human CD2	
13	19	70.4	1514	2	AAV63445		Aav63445 Human CD2	
14	19	70.4	1514	2	AAV81203		Aav81203 Human CD2	
15	19	70.4	1514	3	AAAS0581		Aaa0581 Human cel	
16	19	70.4	1514	3	AAZZ2325		Aazz2325 Human CD2	
17	19	70.4	1514	4	AAS03175		Aas03175 Human lym	
18	19	70.4	1514	6	ABA99035		Abas9035 Human CD2	
19	19	70.4	1514	10	ADD5537		Add5537 Binding d	
20	19	70.4	1514	11	ADI31990		Adi31990 Human cdN	
21	19	70.4	1514	12	ADO49343		Ado49343 Human CD2	

ALIGNMENTS

RESULT 1
AAA10270
ID AAA10270 standard; DNA; 27 BP.
XX
AC
AAA10270;
XX
DT 03-JUL-2000 (first entry)
XX
DE Human CD28 fragment upstream PCR primer.
XX
CD28; cytoplasmic domain; prostate-specific membrane antigen; PSMA;
KW JS91; hybridoma; monoclonal antibody; single chain variable region; scFv;
single chain antibody; fusion receptor; immune response; prostate cancer;
PCR primer; ss.
XX Homo sapiens.
OS
XX
PN WO200014257-A1.
XX
PD 16-MAR-2000.
XX
PP 03-SEP-1999;
PP 99WO-US020349.
XX
PR 04-SEP-1998;
XX
PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX
PI Sadelain M, Bander NH, Gong M;
XX
DR WPI; 2000-257002/22.
XX
PT A fusion receptor composition having the structure:prostate-specific
membrane antigen-single chain variable fragment:optional
connector:cytoplasmic domain, useful for treatment of cancer.
XX
PS Example 6; Page 15; 25pp; English.
XX
CC The invention relates to a novel fusion receptor composition having the
structure: PSMA (prostate-specific membrane antigen)-scFv (single chain
variable fragment):optional connector:cytoplasmic domain, where the
fusion receptor is effective when expressed in a T-cell to promote a
cellular immune response to PSMA. The PSMA-scFv is a single-chain
antibody cloned from the V region genes of hybridoma specific for PSMA,
such as J591. The optional connector is provided to give a spacing

CC between the PSMA-scFv and the cytoplasmic domain, such that both retain CC substantial function. The cytoplasmic domain directs the function of the CC fusion receptor and is generally the cytoplasmic domain of a molecule CC which functions as a transducer of a mammalian immune response in the CC presence of an MHC (major histocompatibility complex)-peptide complex or CC co-stimulatory factor. Examples of cytoplasmic domains that may be CC employed in the present invention include the T-cell receptor gamma-chain CC cytoplasmic domain and the CD28 cytoplasmic domain. In a method of the CC invention, an expression vector encoding the fusion receptor is CC transduced into primary T-lymphocytes obtained from the patient to be CC treated. The transduced lymphocytes are returned to the patient where CC they secrete interleukin-2 (IL-2) and proliferate in response to PSMA- CC positive cells. The resulting cytotoxic lymphocytes specifically lyse CC cells expressing PSMA and can thus be used to target PSMA-positive tumour CC cells. The fusion receptor promotes a cellular immune response to PSMA CC and is useful for the treatment of prostate cancer and other cancers that CC express PSMA. Sequences AAA0270-A10271 represent PCR primers used in an CC exemplification of the present invention to amplify cDNA encoding the CC cytoplasmic and transmembrane domains and part of the extracellular CC domain of human CD28. This was used to construct a gene encoding a fusion CC receptor comprising the PSMA-scFv and the CD28 fragment CC

XX Sequence 27 BP; 6 A; 6 C; 7 G; 8 T; 0 U; 0 Other;

Query Match 100.0% Score 27; DB 3; Length 27;

Best Local Similarity 100.0% Pred. No. 0.003; Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCCGCAATTGAAGTTATGTATCCT 27

Db 1 GCGGCCGCAATTGAAGTTATGTATCCT 27

RESULT 2
ABA92024
ID ABA92024 standard; DNA; 27 BP.

XX AC ABA92024;
XX DT 23-MAY-2002 (First entry)

XX DE Human CD28 cDNA upstream PCR primer.

KW CD28; T cell; receptor; human; antitumour; immunostimulant; cancer; KW therapy; neuroblastoma; melanoma; sarcoma; small lung carcinoma; KW brain tumour; disialoganglioside GD2; imaging; PCR; primer; ss.
XX OS Homo sapiens.
PN US2002018783-A1.
XX PD 14-FEB-2002.

XX PF 30-SEP-1997; 97US-00940544.

XX PR 20-MAR-1997; 97WO-US004427.

XX PA (SADE/) SADELAIN M.
PA (CHBU/) CHEUNG N. V.
PA (KRAU/) KRAUSE A.
PA (GUOH/) GUO H.
XX PI Sadelain M., Cheung NV., Krause A., Guo H.;
DR WPI; 2002-23251/29.

XX New fusion protein, useful for inducing host immune response, comprises PT variable region of light chain of an antibody linked to variable region PT of antibody, CD28 receptor signaling domain and transmembrane domain.
XX Example 3; Page 4; 9pp; English.
XX The present sequence is that of an upstream primer, used with the

CC downstream primer given in ABA92025, in the PCR amplification of a CC segment of human T cell surface receptor CD28 cDNA. The amplified segment CC encodes part of the extracellular domain of CD28, and the transmembrane CC and the cytoplasmic domains. Plasmid pbsecD28 was used as template. The 5' CC primer contains an NotI site, while the 3' primer contains a BamHI site, CC and these sites in the PCR product were used to facilitate insertion into CC retroviral vector SFG. A cell-surface molecule capable of CD28 signalling CC in T cells interacting with disialoganglioside GD2 tumours was CC constructed. This comprised the antigen-binding site of a GP2-specific CC antibody and the transmembrane and signalling domains of the invention CC molecule. This is an example of fusion proteins of the invention comprising a single chain Fv antibody linked to CD28 receptor. T cells CC expressing such fusion proteins exhibit enhanced survival when CC reintroduced to an in vivo environment. They can be used to induce an CC immune response to cells, particularly tumour cells which express the CC antigen to which the antibody is specific. Cells expressing a fusion CC protein including an anti-GD2 scFv are useful for treatment of melanomas, CC neuroblastomas, small lung carcinoma, sarcomas and brain tumours that CC express GD2 as a surface antigen. Cells expressing the fusion proteins of CC the invention can also be used for in vitro purging of stem cells or bone CC marrow and for in vivo targeting of tumour cells and other antigen- CC bearing cells for imaging

XX SQ Sequence 27 BP; 6 A; 6 C; 7 G; 8 T; 0 U; 0 Other;
XX SQ Query Match 100.0% Score 27; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.003; Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCCGCAATTGAAGTTATGTATCCT 27
Db 1 GCGGCCGCAATTGAAGTTATGTATCCT 27

RESULT 3
ADL67231
ID ADL67231 standard; DNA; 328 BP.

XX AC ADL67231;
XX DT 20-MAY-2004 (First entry)
XX DE Human CD28 costimulatory signalling element DNA.

XX KW T cell receptor; TCR; CD28 zeta chain; co-stimulatory signalling region; KW binding element; immunostimulant; therapy; cancer; human; ds.
XX OS Homo sapiens.
PN US2004043401-A1.
XX PR 04-MAR-2004.

XX PF 28-MAY-2003; 2003US-00448256.
XX PR 28-MAY-2002; 2002US-0383872P.
XX PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX PI Sadelain M., Brentjens R., Maher J.;
XX DR WBI; 2004-225696/21.

XX PT New nucleic acid polymer encoding a chimeric T cell receptor having a PT zeta chain portion, useful for treating disorders where the immune PT response needs to be induced, such as cancer.

XX Example 7; SBQ ID NO 6; 25pp; English.
XX The invention relates to a nucleic acid polymer encoding a chimeric T CC cell receptor (TCR) which comprises human CD3 zeta chain intracellular CC domain, a co-stimulatory signalling region and a binding element that CC specifically interacts with a selected target. The methods and

compositions of the invention are useful for treating disorders where the immune response needs to be induced, such as cancer. The present sequence is human CD28 co-stimulatory signalling element DNA.

XX Sequence 328 BP; 75 A; 94 C; 77 G; 82 T; 0 U; 0 Other;
 SQ Query Match 70.4%; Score 19; DB 12; Length 328;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 9 AATTGAACTTATGATATCCT 27
 Db 4 AATTGAACTTATGATATCCT 22

RESULT 4
 AAT96357
 ID AAT96357 standard; cDNA; 660 BP.
 XX AC AAT96357;
 XX DT 08-APR-1998 (first entry)
 XX DE CD28 cDNA.
 XX KW Screening; inhibitor; enhancer; binding; CD28; B7-1; ds.
 XX OS Homo sapiens.
 XX Key Location/Qualifiers
 FH 1..660
 FT /*tag= a
 PT /product= "CD28"
 XX PN EP795554-A2.
 XX PD 17-SEP-1997.
 XX PR 02-OCT-1996; 96JP-00047795.
 XX PP 04-MAR-1997; 97EP-00301438.
 XX PR 05-MAR-1996; 96JP-00047795.
 XX PR 02-OCT-1996; 96JP-00262085.
 XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX PI Hida T, Hattori M, Kurokawa T, Nakaniishi A;
 XX WPI; 1997-450803/42.
 XX DR AAW38413.
 XX PT New xanthene derivatives useful as immunomodulators - e.g., methyl 2-(carboxymethylsulphonyl)-5,7-dichloro-3,8-dihydroxy-6-methyl-9-oxo-9H-xanthene-1-carboxylate.
 XX PS Disclosure; Page 81; 117pp; English.

XX The present sequence was used in the development of a novel method for screening for compounds that inhibit or enhance binding of CD28 to B7-1.
 XX Sequence 660 BP; 164 A; 168 C; 150 G; 178 T; 0 U; 0 Other;
 SQ Query Match 70.4%; Score 19; DB 2; Length 660;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATTGAACTTATGATATCCT 27
 Db 339 AATTGAACTTATGATATCCT 357
 RESULT 5
 ADA02992 standard; cDNA; 663 BP.
 XX AC ADB72730;
 XX ID ADB72730 standard; cDNA; 663 BP.
 XX DT 04-DEC-2003 (first entry)

XX XX XX ADA02992;
 XX XX XX 06-NOV-2003 (first entry)
 DT Human CD28 carcinoma associated coding sequence, SEQ ID NO:1510.
 DE XX
 KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast; prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening; gene; ss.
 XX
 OS Homo sapiens.
 PN WO2003057146-A2.
 XX
 PD 17-JUL-2003.
 XX PF 26-DEC-2002; 2002WO-US041414.
 XX PR 26-DEC-2001; 2001US-00035832.
 XX PA (SAGR-) SAGRES DISCOVERY.
 XX
 P1 Morris DW;
 XX DR WPI; 2003-587068/55.
 XX PT New recombinant nucleic acid encoding carcinoma associated protein, useful for preparing compositions for treating carcinomas.
 XX Claim 1; SEQ ID NO 1510; 245pp; English.
 XX
 CC The invention relates to recombinant carcinoma associated (CA) nucleic acid sequences from mouse and human (ADA01482/ADA03094), and to recombinant carcinoma associated proteins (CAP) encoded by them. The invention also encompasses expression vectors and host cells comprising a CA nucleic acid, a polypeptide (especially an antibody) that specifically binds to the protein, and a biochip comprising CA nucleic acid or fragments thereof. The sequences of the invention were identified using oncogenic retroviruses, which insert into the genome of the host organism at random. Many of these do not carry transduced host oncogenes or pathogenic trans-acting viral genes, meaning that cancer incidence is a direct consequence of the effects of proviral integration into host protooncogenes. The CA nucleic acid sequences can be used to diagnose carcinoma (especially breast cancer, prostate cancer, lymphoma or leukaemia) or a propensity to carcinoma by determination of the sequence of a CA gene, or by determination of CA gene expression in particular tissues. CA nucleic acids, proteins and antibodies are also useful as therapeutic agents and in screening and evaluating drug candidates. The present sequence represents a specifically claimed human CA nucleic acid sequence of the invention. Note: The complete sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 663 BP; 165 A; 168 C; 151 G; 179 T; 0 U; 0 Other;
 CC Query Match 70.4%; Score 19; DB 9; Length 663;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC Qy 9 AATTGAACTTATGATATCCT 27
 CC ID ADB72730
 CC AC ADB72730 standard; cDNA; 663 BP.
 CC XX
 CC DT 04-DEC-2003 (first entry)

XX Human CD28 cDNA.
 XX human; ds; cytosstatic; gene therapy; vaccine; carcinoma; lymphomas;
 XX cancer; neoplasm; adenocarcinoma; sarcoma; gene.
 XX Homo sapiens.
 PN WO2003008583-A2.
 XX 30-JAN-2003.
 XX 26-DEC-2001; 2001WO-US051291.
 PF XX
 PR 02-MAR-2001; 2001US-00798586.
 PR 23-OCT-2001; 2001US-00004113.
 PR 08-NOV-2001; 2001US-00054802.
 PR 30-NOV-2001; 2001US-0097722.
 PR 20-DEC-2001; 2001US-00034650.
 XX PA (SAGR-) SAGRES DISCOVERY.
 XX Morris DW, Engelhard EK;
 XX WPI; 2003-513603/48.
 XX New recombinant nucleic acid comprising a nucleotide sequence of any of
 PT the carcinoma-associated (CA) genes, useful for screening for drug
 PT candidates for diagnosing or treating carcinomas.
 XX
 XX Claim 1; SEQ ID NO 258; 983pp; English.
 XX PS
 XX The invention relates to a recombinant nucleic acid comprising a
 CC nucleotide sequence selected from any of the fully defined carcinoma-
 CC associated (CA) genes from the 50 tables given in the specification. The
 CC CA proteins are secreted, transmembrane or intracellular proteins. The
 CC recombinant nucleic acids are useful for screening for drug candidates
 CC for diagnosing or treating carcinomas. Sequences given in ADC85215-
 CC ADC85514 represent CA genes of the invention.
 XX SQ Sequence 663 BP; 165 A; 151 G; 179 T; 0 U; 0 Other;
 XX Query Match 70.4%; Score 19; DB 10; Length 663;
 XX Best Local Similarity 100.0%; Pred. No. 42;
 XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 9 AATTGAACTTATGTATCC 27
 Db 339 AATTGAACTTATGTATCC 357
 RESULT 8
 ADN74587
 ID XX Human carcinoma associated (CA) nucleic acid #129.
 ID ADM74587 standard; DNA; 663 BP.
 XX AC 01-JUL-2004 (first entry)
 XX DT Human carcinoma associated (CA) nucleic acid #129.
 XX KW Human; carcinoma associated nucleic acid; CA nucleic acid; gene; ds;
 XX KW carcinoma associated protein; CAP; carcinoma; leukaemia; lymphoma;
 XX KW cytostatic.
 XX OS Homo sapiens.
 XX PN US2004072154-A1.
 XX PD 15-APR-2004.
 XX PF 30-NOV-2001; 2001US-00997722.
 XX PR 22-DEC-2000; 2000US-00747377.
 XX PR 02-MAR-2001; 2001US-00798586.
 XX PA (MORRIS D W,
 XX (ENGE) ENGELHARD E K.
 XX PI Morris DW, Engelhard EK;
 XX DR 2004-328562/30.
 XX PS New carcinoma associated gene or protein, useful for preparing a
 XX PT composition for diagnosing or treating carcinoma e.g., leukemia or
 XX lymphoma.
 XX Claim 1; SEQ ID NO 258; 29pp; English.
 XX
 XX The invention relates to new recombinant nucleic acids. The invention
 CC also relates to a host cell comprising a recombinant nucleic acid or
 CC expression vector, an expression vector comprising a recombinant nucleic
 CC acid, a recombinant protein, a method of screening for drug candidates, a

method of screening for a bioactive agent capable of binding to a carcinoma associated protein (CAP) encoded by a nucleotide sequence, a method of screening for a bioactive agent capable of modulating the activity of a CAP, a method of evaluating the effect of a candidate carcinoma drug, a method of diagnosing carcinoma, a method for inhibiting the activity of a CAP, method of treating carcinomas, a method of neutralising the effect of a CAP and a method of diagnosing carcinoma or propensity to carcinoma. A method of evaluating the effect of a candidate carcinoma drug comprises administering the drug to a patient, removing a cell sample from the patient and determining alterations in the expression or activation of a gene comprising the nucleotide sequence. A method of diagnosing carcinoma comprises determining the expression of one or more genes comprising the nucleic acid sequence in a first tissue type of a first individual and comparing the expression of the gene from a second normal tissue type from the first individual or an unaffected individual, where a carcinoma. A method of inhibiting the activity of a CAP comprises binding an inhibitor to the CAP. Treating carcinomas comprises administering to a patient an inhibitor of CAP.

Neutralising the effect of a CAP comprises contacting an agent specific for the CAP. The polypeptide specifically binds to the protein encoded by the nucleic acid. It comprises an antibody that specifically binds to the protein encoded by the nucleic acid. The nucleic acids are useful for preparing a composition for diagnosing or treating carcinoma e.g., leukaemia or lymphoma. This sequence represents a human carcinoma associated (CA) nucleic acid of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

Sequence 663 BP; 165 A; 168 C; 151 G; 179 T; 0 U; 0 Other;

Query Match 70.4%; Score 19; DB 12; Length 663;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATTGAAAGTTATGTATCC 27
Db 339 AATTGAAAGTTATGTATCC 357

RESULT 9
AAQ28837
ID AAQ28837 standard; cDNA; 762 BP.

XX AAQ28837;
AC XX
DT 25-MAR-2003 (revised)
XX 25-FEB-1993 (first entry)
DE Sequence of the CD28 gene.
XX
KW CD28 protein; T-cell immune response; Tp44; differentiation antigen;
KW membrane protein; ds.
XX
Homo sapiens.
XX
FH Key Location/Qualifiers
FT sig_peptide 100..153 /*tag= a
FT mat_peptide 154..762 /*tag= b
FT misc_feature 538..552 /*tag= c
FT /label= primer M1T
FT misc_feature 556..570 /*tag= d
FT /label= primer C1T

XX
PN W0921561-A1.
XX
PD 13-JUL-1990;
XX
PA (GEHO) GEN HOSPITAL CORP.

PF 09-MAR-1992; 92WO-US001867.

XX PR 08-MAR-1991; 91US-00666809.

XX PA (CYTO-) CYTOMED INC.

XX PI Ko J, IP SH;

XX DR WPI; 1992-331717/40.
DR P-PSDB; AAU27103.

XX Soluble proteins binding to B7 proteins and block antigen presenting cells - are useful in treating T-cell mediated immunosuppression diseases e.g. transplant rejection, auto immune diseases etc.

XX PS Disclosure; Fig 1; 43pp; English.

XX The cDNA sequence in AAQ28837 is derived from Aruffo A and Seed B, PNAS USA 84, 8573-77, 1987. The CD28 cDNA encodes a protein of 220 AAs (AAU27103). It contains a signal peptide from AAs 1-18 which is cleaved during the maturation of CD28 protein. The transmembrane region is AAs 153-179. For the production of soluble CD28 protein, the transmembrane region is deleted. The CD28 protein derived from p9-M1T comprises residues 19-157 and is not soluble; the CD28 protein derived from p9-C1T comprises residues 19-157 and is soluble. M1T and C1T are primers. T7 primer is used as the 5' primer to pair with any 3' primer, e.g. M1T and C1T, in PCR reaction for the procurement of truncated CD28 gene which would direct the production of soluble CD28 protein. DNA encoding AAs 19-151 of CD28 is claimed. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 762 BP; 183 A; 198 C; 200 T; 0 U; 0 Other;

SQ Query Match 70.4%; Score 19; DB 2; Length 762;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATTGAAAGTTATGTATCC 27

Db 438 AATTGAAAGTTATGTATCC 456

RESULT 10
AAQ21167

ID AAQ21167 standard; DNA; 1514 BP.
XX AC AAQ21167;

XX DT 25-MAR-2003 (revised)
XX DT 21-MAY-1992 (first entry)

XX DS Human CD28 antigen coding sequence.
XX KW cloning technique; cell surface antigen; immunodiagnosis; tumour; ss.
XX OS Homo sapiens.

XX FH Key Location/Qualifiers
FT CDS /*tag= a
FT FT /*tag= "CD28"
FT sig_peptide 100..153 /*product= "CD28"
FT FT 100..153 /*tag= b
XX
W0921049-A.
XX PD 23-JAN-1992.
XX PP 13-JUL-1990;
XX
PA 90US-00553759.

XX PR 13-JUL-1990;
XX PA (GEHO) GEN HOSPITAL CORP.
XX
PN W0921561-A1.
XX
PD 17-SEP-1992.

PI Seed B, Aruffo A, Amiot M;
 XX WPI; 1992-056664/07.
 DR P-PSDB; AAR20805.

XX New CD53 cell surface antigen and DNA encoding it - for immuno-therapy and diagnosis of haematopoietic neoplasms, etc.

PS Example 3; Fig 7; 160PP; English.

XX A library of recombinants having inserts greater than 0.8kb in size was prepared in the pH3M vector (see AAQ01166) from 1 microgram of polyA (plus) RNA isolated from the human lymphoblastoid cell line JY. The library was screened for CD28 cDNA clones which were isolated using the antibody enrichment method (see e.g. AAQ21164 or AAQ21165). After the third transfaction, COS cells were panned with a specific anti-CD28 Ab. A Hirt supernatant was generated and transformed into E.coli. DNA was prepared from the resulting colonies and transfected into COS cells. Surface expression of CD28 antigen was detected in 3 of the 8 transfected cultures by indirect immunofluorescence. The cDNA insert from one of the positive clones was sequenced. The cDNA insert on 25-MAR-2003 to correct PA field.)

SQ Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;
 XX Query Match 70.4%; Score 19; DB 2; Length 1514;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATTGAAGTTATGTATCCT 27
 DB 438 AATTGAAGTTATGTATCCT 456

RESULT 12
 AAT14707 standard; cDNA; 1514 BP.
 XX ID AAT14707
 AC AAT14707;
 XX DT 25-MAR-2003 (revised)
 ID 30-OCT-1996 (first entry)
 XX DE Human CD28 cDNA.
 XX KW Cell surface antigen; cloning; immunoselection; immunotherapy; therapy;
 KW diagnosis; vector; piH3M; CD28; COS; T-lymphocyte; ss.
 XX OS Homo sapiens.
 XX Key Location/Qualifiers
 FH 100 . 762
 FT /*tag= a
 XX PR 18-OCT-1993;
 PN WO9624380-A1.
 XX PR 25-FEB-1988;
 PD 15-AUG-1996.
 XX PR 88US-00160416.
 XX PR 13-JUL-1989;
 PR 05-FEB-1996; 96WO-US001507.
 XX PR 90US-00379076.
 PR 09-FEB-1995; 95US-00387041.
 PR 18-SEP-1995; 95US-00529878.
 PA (ICNC) ICN PHARM INC.
 XX PR PI 01-DEC-1992;
 XX PR 90US-00553759.
 XX PR 92US-00983647.
 XX PA (GEHO) GEN HOSPITAL CORP.
 XX PR Seed B, Aruffo A;
 XX DR WPI; 1996-200279/20.
 XX DR P-PSDB; AAR1433.
 XX PR Cloning of cDNA encoding cell surface antigen - useful for isolation of diagnostic and therapeutic proteins.

PS Example 3; Fig 7A-B; 79pp; English.
 XX This nucleotide sequence comprises human CD28 cDNA. The cDNA was isolated
 CC from HPA-ALL human T-cell tumour cells using a novel method for cloning
 CC cDNAs from mammalian expression libraries. The method is based on
 CC selection of cells expressing the antigen by adhesion to an antibody-
 CC coated substrate. The method is useful for the isolation and molecular
 CC cloning of any protein which can be expressed and transported to the cell
 CC surface membrane of a eukaryotic cell. It has been used to clone genes
 CC (see AAV81198-220) encoding cell surface antigens such as CD1a, CD1b,
 CC CD1c, CD6, CD7, CD13, CD14, CD16, CD19, CD20, CD22, CD26,
 CC CD27, CD28, CD31, CDw32B, CD32B, CD33, CD34, CD36, CD37,
 CC CD40, CD43, CD53, ICAM-1, FcR_{IIb}, TfR_{II} and Leu8 (see AAW86188
 CC -62, AAB98151-52 and AAW88451). CD40 cDNA (see AAV81198) is specifically
 CC claimed. CD28 polypeptide (see AAW88451) has been expressed in
 CC transfected COS cells.

XX Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;
 SQ Query Match 9 AATGAAAGTTATGATCCCT 27
 Best Local Similarity 100.0%; Prod. No. 47;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATGAAAGTTATGATCCCT 27
 Db 438 AATGAAAGTTATGATCCCT 456

RESULT 15
 AAA50581 ID AAA50581 standard; cDNA; 1514 BP.
 XX DT 19-DEC-2000 (first entry)
 XX DE Human cell surface antigen CD28 cDNA.
 XX KW CD28; cell surface antigen; human; immunoselection; panning;
 KW immunodiagnosis; diagnosis; immunotherapy; gene therapy; immune disorder;
 KW infection; asthma; immune-complex disease; amyloidosis;
 KW multiple sclerosis; parasitic disease; ss.

XX OS Homo sapiens.
 XX PH Key Location/Qualifiers
 FT CDS 100. .762
 FT /*tag= a
 FT sig_peptide 100. .153
 FT /*tag= b
 FT mat_peptide 154. .759
 FT /*tag= c

XX PN US611093-A.
 XX PD 29-AUG-2000.

XX PF 28-OCT-1998; 98US-00161612.
 XX PR 25-FEB-1998; 88US-00160416.
 PR 13-JUL-1998; 89US-0037006.
 PR 23-MAR-1990; 90US-00498809.
 PR 13-JUL-1990; 90US-00553759.
 PR 01-DEC-1992; 92US-00983647.
 XX PA (GEHO) GEN HOSPITAL CORP.
 XX PI Stamenkovic I, Seed B;
 XX DR WPI; 2000-586382/55.
 DR P-PSDB; AAY96128.

PS Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
 PT useful for immunodiagnosis and immunotherapy of immune-mediated
 PT infections or disorders, e.g. asthma, immune-complex disease, parasitic
 PT diseases.

XX Example 3; Fig 7A-B; 75pp; English.

XX The present sequence is that of cDNA encoding human cell surface antigen
 CC (CSA) CD28 (see AAY6128). The cDNA was derived from a human T-cell
 CC tumour line HPA-ALL cDNA library using a new method for cloning CSA
 CC cDNAs. The method is based upon transient expression of CSA in eukaryotic
 CC cells and physical selection of cells expressing the antigen by adhesion
 CC to (panning on) an antibody-coated substrate such as a culture dish. The
 CC predicted amino acid sequence of CR28 suggests an integral membrane
 CC protein with a single membrane-spanning hydrophobic domain terminating in
 CC a 41-amino acid cytoplasmic domain. CSA nucleic acids isolated by the
 CC method of the invention, and the proteins they encode, are useful for
 CC immunodiagnostic and immunotherapeutic applications, including the
 CC diagnosis and treatment of immune-mediated infections, diseases, and
 CC disorders in animals, including humans. These disorders include asthma,
 CC immune-complex disease, amyloidosis, parasitic diseases or multiple
 CC sclerosis.

XX Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;
 SQ Query Match 9 AATGAAAGTTATGATCCCT 27
 Best Local Similarity 100.0%; Prod. No. 47;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATGAAAGTTATGATCCCT 27
 Db 438 AATGAAAGTTATGATCCCT 456

RESULT 16
 AAZ29325 ID AAZ29325 standard; cDNA; 1514 BP.
 XX DT 29-FEB-2000 (first entry)
 XX DE Human CD28 receptor cDNA.
 XX KW Human CD28; B7 molecule; immune response; cell surface receptor;
 KW Major histocompatibility complex; MHC ClassII; proton motor force;
 KW mitochondrial membrane potential; mitochondrial metabolism; cancer;
 KW autoimmune disease; neurodegenerative disorder; ds.
 XX OS Homo sapiens.
 XX PH Key Location/Qualifiers
 FT CDS 100. .762
 FT /*tag= a
 FT sig_peptide 100. .153
 FT /*tag= b
 FT mat_peptide 154. .759
 FT /*tag= c
 XX PN WO953953-A2.
 XX PD 28-OCT-1999.
 XX PR 30-MAR-1999;
 PR 17-APR-1998;
 PR 29-JUL-1998;
 PR 24-SEP-1998;
 XX PA (UYVE-) UNIV VERMONT.
 XX PI Newell MK;
 XX DR WPI; 2000-09673/08.
 DR P-PSDB; AAY44294.
 XX

PT Use of cell surface and membrane characteristics for developing products
for treating cancers, autoimmune diseases or neurodegenerative diseases.

XX Disclosure; Page 120-121; 123pp; English.

CC The present sequence encodes human CD28 receptor. CD28 is a homodimeric
CC glycoprotein involved in a secondary signalling pathway in the activation
CC of T-cell proliferation. B7 on nerve cells can interact with CD28 on the
CC immune cell leading to immune cell activation. The regulation of cell
CC expression of MHC classII and co-stimulatory molecule B7 can be
CC manipulated by regulating the intracellular dissipation of proton motor
CC force which can be assessed in terms of mitochondrial membrane potential.
CC These methods can be used for regulating cell growth and division to
CC control disease processes by manipulating mitochondrial metabolism and
CC the expression of cell surface immune proteins. They can be used for
CC treating diseases associated with excessive cellular division, aberrant
CC differentiation, and premature cellular death, e.g. cancers, autoimmune
CC diseases, neurodegenerative disorders etc

XX Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;

SQ Query Match 70.4%; Score 19; DB 3; Length 1514;

XX Best Local Similarity 100.0%; Pred. No. 47;

XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 9 AATTGAAGTTATGATCTT 27

Db 438 AATTGAAGTTATGATCTT 456

RESULT 17
AAS03175 ID AAS03175 standard; cDNA; 1514 BP.

AC XX
XX DT 29-AUG-2001 (first entry)

DB Human lymphocyte cell surface antigen CD28 cDNA sequence.

XX Human; lymphocyte cell surface antigen; immune-mediated disease; CD28;

XX infection; immune deficiency disorder; hypersensitivity; inflammation;

XX systemic lupus erythematosus; platelet disorder; rheumatoid arthritis;

XX transplant rejection; asthma; ss.

XX Homo sapiens.

PH Key Location/Qualifiers
PT CDS /*tag= a

PT /*product= "CD28 antigen"

PT big_peptide /*tag= b

PT mat_peptide 154..759 /*tag= c

XX US6218525-B1.

XX PD 17-APR-2001.

XX PP 01-DEC-1992; 92US-00983647.

XX PR 25-FEB-1988; 88US-00160416.

PR 13-JUL-1989; 89US-00379076.

PR 13-JUL-1990; 90US-00533755.

XX (GBHO) GEN HOSPITAL CORP.

XX Seed B, Aruffo A, Simmons D;

XX DR WPI; 2001-289848/30.

DR P-PSDB; AAU02437.

XX DR WPI; 2002-187719/24.

PT New recombinant DNA encoding CD28 useful for diagnosing and treating
PT immune-mediated diseases, infections or disorders, e.g. systemic lupus
PT erythematosus, asthma, transplant rejection, rheumatoid arthritis.

PS Example 3: Fig 7A-7B; 72pp; English.

XX The present sequence encoding for human lymphocyte cell surface antigen

CC CD28 is used to obtain a new genetically engineered cDNA sequence 1
CC encoding the CD28 amino acid extracellular domain sequence (amino acids 1
CC to 134 given in AAU02437) and/or comprising nucleotides 100..755, 154..555 or
CC 154..759 of the CD28 cDNA sequence. Various human lymphocyte cell surface

CC antigen cDNA sequences (AAS03172, AAS03173, AAS03175-AAS03195) are
CC described in the present invention. The invention relates to a novel
CC method of cloning cDNA encoding cell surface antigens and efficient
CC construction of cDNA libraries. Also described are 2 expression vectors

CC (AAS03171, AAS03174) which provide high level expression in eukaryotic
CC host cells. The purified genes and proteins are useful for
CC immunodiagnostic and immunotherapeutic applications, such as in the
CC diagnosis and treatment of immune-mediated diseases, infections or

CC disorders in animals and humans. Such diseases include immune deficiency
CC diseases, diseases of immediate type of hypersensitivity, asthma,
CC hypersensitivity pneumonitis, systemic lupus erythematosus, rheumatoid
CC arthritis, acute and chronic inflammation, platelet disorders, plasma and
CC other cell neoplasms, parasitic diseases, multiple sclerosis, Guillain-
CC Barre Syndrome and tissue and organ transplant rejection. The sequences
CC can also be used to identify, isolate and purify other antibodies and
CC antigens

XX SQ Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;

XX SQ Query Match 70.4%; Score 19; DB 4; Length 1514;
XX Best Local Similarity 100.0%; Pred. No. 47;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 9 AATTGAAGTTATGATCTT 27

Db 438 AATTGAAGTTATGATCTT 456

RESULT 18
ABA99035 ID ABA99035 standard; DNA; 1514 BP.

AC XX
XX AC ABA99035;

XX DT 13-JUN-2002 (first entry)

XX Human CD28 gene.

XX Human; CD28; immune response; T cell; graft-versus-host disease; GVHD;

XX marrow transplant rejection; organ transplant rejection; psoriasis;

XX tissue transplant rejection; autoimmune disease; diabetes mellitus;

XX multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;

XX systemic sclerosis; dermatomyositis; Sjogren's syndrome;

XX polyarteritis nodosa; vasculitis; sepsis; gene; ds.

XX OS Homo sapiens.

XX PN US2002006403-A1.

XX PD 17-JAN-2002.

XX PR 14-DEC-2000; 20000US-00738546.

XX PR 14-DEC-1999; 99US-0170857P.

XX PA (YUXX/) YU X.

PA (ANAS/) ANASETTI C.

XX PI Yu X, Anasetti C;

XX DR WPI; 2002-187719/24.

XX DR WPI; 2002-187719/24.

PT	subject having or suspected of having a malignant condition or a B-cell disorder, e.g. melanoma, Grave's disease or autoimmune disease.
PT	rheumatoid arthritis, polyarthritis, by administering anti-CD28 antibody.
XX	
PS	Disclosure; Page 28; 37pp; English.
XX	
CC	The sequence represents the human CD28 gene. The invention relates to a novel method for inhibiting an immune response by administering an anti-CD28 antibody. The method of the invention has immunosuppressive, antipsoriatic, antiarthritic, antiarthritis, antirheumatic, neuroprotective, dermatological, vasoconstrictive, antiflammatory, antianesthetic, hepatotropic, antiulcer, antibacterial and ophthalmological activity. The method works to inhibit immune response by reversing or blocking T cell activation. The method is useful for inhibiting an immune response in a subject susceptible to graft-versus-host disease (GVHD), marrow transplant rejection, or having autoimmune disease including psoriasis, diabetes mellitus, multiple sclerosis, rheumatoid arthritis, systemic lupus erythematosus, systemic sclerosis, dermatomyositis, polymyositis, Sjogren's syndrome, vasculitis, nodosa or vasculitis. The method is also useful for treating sepsis, and other autoimmune diseases including autoimmune hepatitis, autoimmune haemolytic anaemia, Behcet's disease, myasthenia gravis, cirrhosis, uveitis, ulcerative colitis, and vitiligo.
XX	
SQ	Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;
Query Match	70.4%; Score 19; DB 6; Length 1514;
Best Local Similarity	100.0%;
Matches	19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	9 AATTGAAGTTATGATCTT 27
Db	438 AATTGAAGTTATGATCTT 456
RESULT 19	
ID	ADD25537 standard; DNA; 1514 BP.
XX	
AC	ADD25537;
XX	15-JAN-2004 (first entry)
DB	Binding domain-immunoglobulin fusion protein-associated DNA #53.
XX	ds; Binding domain; immunoglobulin; fusion protein; cytosolic; antiarthritic; immunosuppressive; antidiabetic; antithyroid; neuroprotective; hinge region; immunoglobulin heavy chain; CH2 constant region; CH3 constant region; IgG1; antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation; malignant condition; B-cell disorder; melanoma; carcinoma; rheumatoid arthritis; myasthenia gravis; Grave's disease; sarcoma; type I diabetes mellitus; multiple sclerosis; autoimmune disease.
XX	Unidentified.
OS	
PN	US2003118592-A1.
XX	
PD	26-JUN-2003.
XX	25-JUL-2002; 2002US-00207655.
XX	17-JAN-2001; 2001US-0367358P.
PR	17-JAN-2002; 2002US-0367358P.
PR	03-JUN-2002; 2002US-0385691P.
PA	(GENE-) GENECRAFT INC.
XX	Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
DR	WPI; 2003-801317/75.
XX	New binding domain-immunoglobulin fusion protein, useful for treating a
PT	subject having or suspected of having a malignant condition or a B-cell disorder, e.g. melanoma, Grave's disease or autoimmune disease.
PT	Disclosure; SEQ ID NO 98; 157pp; English.
XX	
CC	Unidentified
SQ	Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;
Query Match	70.4%; Score 19; DB 10; Length 1514;
Best Local Similarity	100.0%;
Matches	19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	9 AATTGAAGTTATGATCTT 27
Db	438 AATTGAAGTTATGATCTT 456
RESULT 20	
ID	AD131990 standard; cDNA; 1514 BP.
XX	
AC	AD131990;
XX	17-JUN-2004 (first entry)
DT	XX
DE	Human cDNA #1316.
KW	Human; gene; SB; immunological response; immunopathological condition; Crohn's disease; asthma; ulcerative colitis; hypereosinophilia; irritable bowel syndrome; osteoarthritis; rheumatoid arthritis; acute monocytic leukemia; antiinflammatory; antiasthmatic; antiulcer; osteopathic; antiarthritic; antirheumatic; cytostatic.
KW	Homo sapiens.
XX	
EN	US6601879-B1.
XX	19-AUG-2003.
PP	US6601879-B1.
XX	09-FEB-1998;
PR	09-FEB-1998;
XX	98US-00023655.
PA	(INCYT-) INCYTE CORP.
XX	Cocks BG, Stuart SG, Seilhamer JJ;
PI	WPI; 2003-895307/82.
XX	A composition comprising a plurality of cDNAs, useful for detecting altered expression of genes in an immunological response or for diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma or osteoarthritis.
XX	Claim 1; SEQ ID NO 1316; 50pp; English.
PS	The invention relates to a composition comprising a plurality of cDNAs for detecting the altered expression of genes in an immunological response. The invention also relates to a method of diagnosing or monitoring the treatment of an immunopathological condition in a sample, comprising obtaining nucleic acids from a sample, contacting the nucleic acids of the sample with an array comprising the plurality of cDNAs under conditions to form one or more hybridisation complexes, detecting the hybridisation complexes and comparing the levels of the detected hybridisation complexes with the level of hybridisation complexes detected in a non-diseased sample, where an altered level of the detected hybridisation complexes correlates with the presence of an immunopathological condition. Also disclosed are an expression profile comprising a microarray and a plurality of detectable complexes and a method for identifying a plurality of polynucleotide probes. The cDNAs are useful as hybridisable array elements in a microarray for monitoring the expression of target polynucleotides. The microarray can be used in

CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,
 CC ulcerative colitis, hyperesinophilia, irritable bowel syndrome,
 CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in
 CC identifying agents for the treatment of the diseases. The microarray may
 CC also be used in drug discovery and development, toxicological and
 CC carcinogenicity studies, forensics or pharmacogenomics. The composition
 CC may also be used in purification of a subpopulation of mRNAs, cDNAs or
 CC genomic fragments. This sequence represents a human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification but was obtained in electronic format directly
 CC from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;
 XX Query Match 70.4%; Score 19; DB 11; Length 1514;
 XX Best Local Similarity 100%; Pred. No. 47;
 XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 9 AATTGAAGTTATGATATCCT 27
 Db 438 AATTGAAGTTATGATATCCT 456

RESULT 21
 ADO49343

ID ADO49343 standard; cDNA; 1514 BP.

XX

AC ADO49343;

XX

DT 15-APR-2004 (first entry)

XX

DE Human CD28 antigen cDNA.

XX

KW cell surface antigen; immune-mediated disorder; asthma; inflammation; ss;
 KW gene; human.

XX

Homo sapiens.

XX

US2004072283-A1.

XX

PD 15-APR-2004.

XX

PP 17-APR-2001; 2001US-00836544.

XX

PR 25-FEB-1988;

PR 13-JUL-1989;

PR 23-MAR-1990;

PR 13-JUL-1990;

PR 01-DEC-1992;

XX

(SEED/) SEED B.

PA (ALLE/) ALLEN J.

PA (ARUFI/) ARUFFO A.

PA (CAME/) CAMERINI D.

PA (LAUFI/) LAUFER L.

PA (OQUE/) OQUENDO C.

PA (STAM/) STAMENKOVIĆ I.

PA (STEN/) STENGELIN S.

PA (AMIO/) AMIOT M.

XX

PA Seed B, Allen J, Aruffo A, Camerini D, Lauffer L, Oquendo C;

PA Simmons D, Stamenkovic I, Stengelin S, Amiot M;

XX WPI; 2004-328571/30.

DR P-PSDB; ADO49344.

XX New cloning cDNA segments encoding cell surface antigens of human lymphocytes, useful in diagnosing and treating asthma, rheumatoid arthritis, multiple sclerosis, vasculitis and inflammation and infections.

XX

PT PT

PT PT

PT PT

PT PT

PT PT

PS Example 3; Fig 7; 75pp; English.

XX The invention relates to a cloned cDNA segment encoding a cell surface CC antigen selected from CD1a, CD1b, CD1c, CD2, CD6, CD7, CD14, CC CD19, CD20, CD22, CD26, CD27, CD31, CD32b, CD33, CD34, CD36, CC CD37, CD38, CD39, CD40, CDw32a, CDw32b, and their functional derivatives. The cell CC surface antigens of human lymphocytes prepared from the cDNAs are useful CC in diagnostic and therapeutic utility in immune-mediated disorders CC (asthma, rheumatoid arthritis, multiple sclerosis, vasculitis and CC inflammation) and infections in mammals, including humans. The present CC sequence represents a human cell surface antigen cDNA.

XX Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;

SQ Query Match 70.4%; Score 19; DB 12; Length 1514;

XX Best Local Similarity 100%; Pred. No. 47;
 XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATTGAAGTTATGATATCCT 27

Db 438 AATTGAAGTTATGATATCCT 456

RESULT 22

AAS89399/C

ID AAS89399 standard; cDNA; 1557 BP.

XX

AC AAS89399;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #25203.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss; Homo sapiens.

OS XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PP 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

XX

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSEQ-) HYSEQ INC.

XX

P1 Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-339362/73.

DR P-PSDB; ABG25212.

XX

XX New isolated polynucleotide and encoded polypeptides, useful in sequencing (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

XX

XX Claim 1; SEQ ID NO 25203; 103pp; English.

XX

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The

XX

CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

CC Sequence 1557 BP; 392 A; 398 C; 361 G; 406 T; 0 U; 0 Other;	XX CC Revised record issued on 09-SEP-2004 : Correction to keywords
Qy XX SQ Sequence 1574 BP; 423 A; 375 C; 347 G; 429 T; 0 U; 0 Other;	
Db XX Qy Query Match 70.4%; Score 19; DB 1; Length 1574;	
XX Best Local Similarity 100.0%; Pred. No. 48;	
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
XX	
Qy 9 AATTGAAGTTATGTATCCCT 27	
Db 193 AATTGAAGTTATGTATCCCT 175	
XX	
RESULT 24	
AAL49546	
ID AAL49546 standard; DNA; 3803 BP.	
XX	
AC AAL49546;	
XX	
DT 27-NOV-2002 (first entry)	
XX	
DE Human CD28 coding sequence.	
XX	
KW Human; CD28; immune response; signaling; diabetes; autoimmune disease;	
KW immunosuppressive; dermatological; antiinflammatory; antithyroid;	
KW antirheumatic; antiarthritic; nootropic; allergy; rheumatoid arthritis;	
KW systemic lupus erythematosus; myasthenia gravis; autoimmune thyroiditis;	
KW vitiligo; alopecia; inflammatory bowel disease; Addison's disease;	
KW Graves disease; haemolytic anaemia; Sjogren's syndrome; gene; ds.	
XX	
OS Homo sapiens.	
XX	
DB CD28 antigen cDNA.	
XX	
KW Tyrosine suppressor gene; ASV LTR; human cytomegalovirus AD169 enhancer;	
KW HIV box; immunoselection; immune deficiency diseases; vasculitis;	
KW systemic lupus erythematosus; rheumatoid arthritis; neoplasms; ss.	
XX	
OS Homo sapiens.	
OS Unidentified.	
XX	
Key Location/Qualifiers	
CDS PH XX 29-AUG-2002.	
AC XX 100. .819	
AC XX 15-FEB-2002; 2002WO-US004772.	
AC XX 100. .274	
AC XX 16-FEB-2001; 2001US-0269756P.	
AC XX (GEMY) GENETICS INST INC.	
AC XX PA	
AC XX O'hara RM, Nagelin AM;	
AC XX XX DR WPI; 2002-674899/72.	
AC XX XX PD WPI; AA019101.	
AC XX PR Downmodulating immune responses by blocking CD28-mediated signaling,	
AC XX PT useful for preventing and/or treating autoimmune disorders such as	
AC XX PT diabetes, allergic reactions, graft versus host disease, systemic lupus	
AC XX PT erythematosus.	
AC XX Disclosure; Page 60-61; 61pp; English.	
AC XX	
AC CC The present invention relates to a method of therapeutically	
AC CC downmodulating an autoimmune response or an ongoing autoimmune response,	
AC CC comprising administering an antigen binding portion of an anti-CD28	
AC CC antibody that blocks signaling via CD28 to the subject so that an	
AC CC autoimmune response or an ongoing autoimmune response in the subject is	
AC CC downmodulated. The methods are useful in therapeutically and	
AC CC prophylactically downmodulating the immune response in subjects having	
AC CC autoimmune disorders such as diabetes, allergy and allergic reactions,	
AC CC transplantation rejection, graft versus host disease, systemic lupus	
AC CC erythematosus, rheumatoid arthritis, myasthenia gravis, autoimmune	
AC CC thyroiditis, vitiligo, alopecia, inflammatory bowel disease, Addison's	
AC CC disease, Graves disease, haemolytic anaemia and Sjogren's syndrome. The	
AC CC present sequence is the human CD28 coding sequence	

composition for diagnosing or treating carcinoma e.g., leukemia or lymphoma.

Claim 1: SEQ ID NO 257: 29pp; English.

The invention relates to new recombinant nucleic acids. The invention also relates to a host cell comprising a recombinant nucleic acid or expression vector, an expression vector comprising a recombinant nucleic acid, a recombinant protein, a method of screening for drug candidates, a method of screening for a bioactive agent capable of binding to a carcinoma associated protein (CAP) encoded by a nucleotide sequence, a method of screening for a bioactive agent capable of modulating the activity of CAP, a method of evaluating the effect of a candidate carcinoma drug, a method of diagnosing carcinoma, a method for inhibiting the activity of a CAP, a method of treating carcinoma, a method of neutralising the effect of a CAP and a method of diagnosing carcinoma or propensity to carcinoma. A method of evaluating the effect of a candidate carcinoma drug comprises administering the drug to a patient, removing a cell sample from the patient and determining alterations in the expression or activation of a gene comprising the nucleotide sequence. A method of diagnosing carcinoma comprises determining the expression of one or more genes comprising the nucleic acid sequence in a first tissue type of a first individual and comparing the expression of the gene from a second normal tissue type from the first individual or a second unaffected individual, where a difference in the expression indicates that the first individual has carcinoma. A method of inhibiting the activity of a CAP comprises binding an inhibitor to the CAP. Treating carcinomas comprises administering an inhibitor of CAP. Neutralising the effect of a CAP comprises contacting an agent specific for the CAP. The polypeptide specifically binds to the protein encoded by the nucleic acid. It comprises an antibody that specifically binds to the protein encoded by the nucleic acid. The nucleic acids are useful for preparing a composition for diagnosing or treating carcinoma e.g., leukaemia or lymphoma. This sequence represents a human carcinoma associated (CA) nucleic acid of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at

卷之三

Human; cytotoxic T lymphocyte antigen-4; CTLA-4; CD28; AIDS; therapy; Acquired immune deficiency syndrome; neuroprotective; dermatological; immune response; organ transplantation; autoimmune disease; allergy; SLE; systemic lupus erythematosus; multiple sclerosis; tumour vaccination; immunodeficiency disease; DiGeorge Syndrome; cancer; ds.

9 ATTTGAAGTATTGATCC 27

561 AATTGAAAGTTATGTATCCT 579

RESULT 30

ENCL:JGZ ADN 05855 standard; CDNA; 3804 BP.

ADDRESS: 8 EEE 1

Antipsoriatic cDNA sequence #1159.
ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.

卷之三

K

MOZU10204 / S-HZ:

08-ABR-2004.

25-SEP-2003 ; 2003WO-US030907.

25-SEP-2002; 2002US-041

GENEVA, SWITZERLAND

Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
 PPI Wu TD;
 XX
 DR WPI: 2004-305105/28.
 XX
 New PRO nucleic acid or polypeptide, useful for preparing a pharmaceutical composition for diagnosing or treating psoriasis in a mammal.
 XX
 PS Claim 1; SEQ ID NO 2250; 3069pp; English.
 XX
 The invention relates to novel polynucleotide and polypeptides for treating psoriasis or a sequence having at least 80% identity to the above sequences. The nucleic acid is useful for preparing a composition for diagnosing or treating psoriasis in a mammal. This sequence corresponds to one of the polynucleotides of the invention.
 XX
 Sequence 3804 BP; 1076 A; 766 C; 880 G; 1082 T; 0 U; 0 Other;
 SQ Query Match 70.4%; Score 19; DB 12; Length 3804;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 YY 9 AATTGAAAGTTATGTATCT 27
 Db 561 AATTGAAAGTTATGTATCT 579
 XX
 RESULT 31
 AAD21976 standard; DNA; 3806 BP.
 ID AAD21976
 XX
 AC AAD21976
 XX
 DT 12-PFB-2002 (first entry)
 XX
 DE Human CD28 DNA.
 XX
 KW Human; cytotoxic T lymphocyte antigen-4; CTLA-4; CD28; AIDS; therapy;
 KW acquired immune deficiency syndrome; neuroprotective; dermatological;
 KW immune response; organ transplantation; autoimmune disease; allergy; SLE;
 KW systemic lupus erythematosus; multiple sclerosis; tumour vaccination;
 KW immunodeficiency disease; DiGeorge Syndrome; cancer; ds.
 XX
 OS Homo sapiens.
 XX
 FH Location/Qualifiers
 FT 223 .885
 FT /*tag= a
 FT /product= "Human CD28 protein"
 XX
 WO200179300-A1.
 XX
 PD 25-OCT-2001.
 XX
 PF 12-APR-2001; 2001WO-US012275.
 XX
 PR 12-APR-2000; 2000US-0196831P.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 PI Bluestone JA, Collins M, Whitters M, Griffin M, Kranz D;
 XX
 DR WPI: 2002-017603/02.
 XX
 P-PSDB; AAE13721.
 XX
 New construct for downmodulating immune response in a subject, has exposed surface attached with antigen-binding portion of antibody that binds to cytotoxic T lymphocyte antigen-4 and major histocompatibility molecule.

PS Disclosure; Page 93-94; 98pp; English.

XX The invention relates to a construct for downmodulating immune response in a subject. The construct comprises an exposed surface attached with an antigen-binding portion of an antibody that binds to a cytotoxic T lymphocytic antigen (CTLA)-4 or CD28 expressed on T-cell of the subject, and a major histocompatibility complex (MHC) molecule. The construct is useful for treating an individual afflicted with a disease or disorder that would benefit from downregulation of immune response. Downmodulation of the immune response is useful to downmodulate the immune response in situations of tissue, skin and organ transplantation, graft-versus-host disease or in autoimmune diseases such as systemic lupus erythematosus and multiple sclerosis. The construct is useful for inhibiting immune cell activation and prevents production of autoantibodies or cytokines which may be involved in disease process. Inhibition of immune cell activation is useful in the treatment of allergy and allergic reactions e.g., by inhibiting IgE production. The construct is also useful for enhancing desirable immune response in a subject e.g., in situations of tumour vaccination, in viral immunity or in immunodeficiency diseases such as AIDS and DiGeorge Syndrome. Induction/enhancement of immune cell function results in increased tumour destruction in cancer patients. The construct is also useful in treating infectious diseases. The present sequence is human CD28 DNA.

Sequence 3806 BP; 1079 A; 765 C; 881 G; 1081 T; 0 U; 0 Other;

XX Sequence 3806 BP; 1079 A; 765 C; 881 G; 1081 T; 0 U; 0 Other;
XX Sequence 3806 BP; 1079 A; 765 C; 881 G; 1081 T; 0 U; 0 Other;

Query Match 70.4%; Score 19; DB 6; Length 3806;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATTGAAAGTTATGTATCC 27
Db 561 AATTGAAAGTTATGTATCC 579

RESULT 32
ADA02990 Standard; DNA; 51365 BP.
XX DT 06-NOV-2003 (first entry)

XX Human CD28 carcinoma associated gene, SEQ ID NO:1508.
XX Human; carcinoma associated; oncogene; breast;
KW prostate; lymphoma; leukaemia; cytosstatic; gene therapy; drug screening;
KW gene; ds.
OS Homo sapiens.
XX WO2003057146-A2.

XX PR 02-MAR-2001; 2001US-00798586.
XX PR 17-JUL-2003.

XX PR 26-DEC-2002; 2002WO-US041414.
XX PR 26-DEC-2001; 2001US-00035832.

XX PA (SAGR-) SAGRES DISCOVERY.
XX Morris DW;
PI Morris DW, Engelhard EK;
XX DR WPI; 2003-239337/23.

XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas, cancers, neoplasm, adenocarcinoma, or sarcomas.
PT PT
PT Claim 1; SEQ ID NO 556; 2304pp; English.
XX PS
XX The invention relates to a novel recombinant nucleic acid comprising a nucleotide sequence selected from any of the 660 sequences fully defined in the specification. A polynucleotide of the invention has cytosstatic activity, and may have a use in gene therapy or in a vaccine. The recombinant carcinoma associated proteins (CAP) encoded by them. The

CC invention also encompasses expression vectors and host cells comprising a CA nucleic acid, a polypeptide (especially an antibody) that specifically binds to the protein, and a biochip comprising CA nucleic acid or fragments thereof. The sequences of the invention were identified using oncogenic retroviruses, which insert into the genome of the host organism at random. Many of these do not carry transduced host oncogenes or pathogenic trans-acting viral genes, meaning that cancer incidence is a direct consequence of the effects of proviral integration into host oncogenes. The CA nucleic acid sequences can be used to diagnose carcinoma (especially breast cancer, prostate cancer, lymphoma or leukaemia) or a propensity to carcinoma by determination of the sequence of a CA gene, or by determination of CA gene expression in particular tissues. CA nucleic acids, proteins and antibodies are also useful as therapeutic agents and in screening and evaluating drug candidates. The present sequence represents a specifically claimed human CA nucleic acid sequence of the invention. Note: The complete sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences.

Sequence 51365 BP; 15176 A; 10059 C; 10474 G; 15656 T; 0 U; 0 Other;

Query Match 70.4%; Score 19; DB 9; Length 51365;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATTGAAAGTTATGTATCC 27
Db 30447 AATTGAAAGTTATGTATCC 30465

RESULT 33
ADB72728
ID ADB72728 standard; DNA; 51365 BP.
XX AC ADB72728;
XX DT 04-DEC-2003 (first entry)
XX DE Human CD28 gene.
XX KW human; ds; cytosstatic; gene therapy; vaccine; carcinoma; lymphomas; cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX OS Homo sapiens.
XX PN WO2003008583-A2.
XX PD 30-JAN-2003.
XX XX 26-DEC-2001; 2001WO-US051291.
XX PR 02-MAR-2001; 2001US-00798586.
XX PR 23-OCT-2001; 2001US-0004113.
XX PR 08-NOV-2001; 2001US-005482.
XX PR 30-NOV-2001; 2001US-0097722.
XX PR 20-DEC-2001; 2001US-00034550.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW, Engelhard EK;
XX DR WPI; 2003-239337/23.

recombinant nucleic acids and polypeptides are useful for treating carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and sarcomas. The present sequence represents a human gene of the invention.

CC Sequence 51365 BP; 15176 A; 10059 C; 10474 G; 15656 T; 0 U; 0 Other;

CC XX SQ

Query Match 70.4%; Score 19; DB 10; Length 51365;

Best Local Similarity 100.0%; Pred. No. 82;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATTGAACTTATGTATCCT 27

Db 30447 AATTGAACTTATGTATCCT 30465

RESULT 34

ADC85410

ID ADC85410 standard; DNA; 51365 BP.

XX

AC ADC85410;

XX

DT 01-JAN-2004 (first entry)

XX

DE Human cd28 genomic sequence.

XX

DB (SAGR-) SAGRES DISCOVERY.

XX

PA Morris DW, Engelhard EK;

XX

PR WPI; 2003-513603/4.

XX

PS Claim 1; SEQ ID NO 256; 983pp; English.

XX

PP 02-DEC-2002; 2002WO-US038582.

XX

PR 30-NOV-2001; 2001US-00997722.

XX

PA (MORRIS D W. (ENGELHARD E K.

XX

QY 9 AATTGAAGTTATGATCCCT 27
 DB 30447 AATTGAAGTTATGATCCCT 30465

RESULT 36
 AAA69169 standard; DNA; 3692 BP.
 XX
 AC AAA69169;
 XX DT 27-OCT-2000 (first entry)
 XX
 DE Bacteriophage Dp-1 nucleotide sequence dp1ORF001.
 XX KW Bacteriophage; antimicrobial; genome; identification; antibacterial;
 bacterial growth inhibition; bacterial infection; ds.
 XX OS Bacteriophage Dp-1.
 XX PN WO20032825-A2.
 XX PD 08-JUN-2000.
 XX PF 08-JUN-2000.
 XX XX 99WO-IB002040.
 XX PR 03-DEC-1999;
 XX PR 03-DEC-1998;
 XX PR 03-JUN-1999;
 XX PR 28-SEP-1999;
 XX PR 30-SEP-1999;
 XX PR 30-SEP-1999;
 XX PR 01-DEC-1999;
 XX PR 02-DEC-1999;
 XX PA (PHAG-) PHAGETECH INC.
 XX PA Pelletier J, Gros P, Dubow M;
 XX DR WPI; 2000-412361/35.

Identifying a bacteriophage coding region for treating bacterial infections comprises identifying a nucleic acid encoding a product that inhibits bacteria when a bacteriophage infects a bacterium.

Example 17; Page 348-358; 456pp; English.

XX CC The present invention describes a method for identifying a bacteriophage coding region encoding a product active on an essential bacterial target. The method comprises identifying a nucleic acid sequence encoding a gene product that provides a bacteria-inhibiting function when an uncharacterised bacteriophage infects a pathogenic bacterium. The compound active on a target of bacteriophage inhibitor protein in a bacteria is used to treat or prevent a bacterial infection in an animal. AAA68243 to AAA69442 and AAB16523 to AAB16954 represent bacteriophage nucleotide and protein sequences which are used in the exemplification of the present invention

XX SQ Sequence 565506 BP; 18170 A; 10773 C; 12019 G; 15544 T; 0 U; 0 Other; DR p-PSDB; AAB16682.

Identifying a bacteriophage coding region for treating bacterial infections comprises identifying a nucleic acid encoding a product that inhibits bacteria when a bacteriophage infects a bacterium.

Example 17; Page 363-364; 456pp; English.

CC The present invention describes a method for identifying a bacteriophage coding region encoding a product active on an essential bacterial target. The method comprises identifying a nucleic acid sequence encoding a gene product that provides a bacteria-inhibiting function when an uncharacterised bacteriophage infects a pathogenic bacterium. The compound active on a target of a bacteriophage inhibitor protein in a bacteria is used to treat or prevent a bacterial infection in an animal. AAA68243 to AAA69442 and AAB16523 to AAB16954 represent bacteriophage nucleotide and protein sequences which are used in the exemplification of the present invention

XX SQ Sequence 3692 BP; 1151 A; 725 C; 842 G; 974 T; 0 U; 0 Other;

Query Match 67.4%; Score 18.2; DB 3; Length 3692;
 Best Local Similarity 87.0%; Prod. No. 1.3e+02;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GCCCCAATTGAAAGTTATGATCC 26
 DB 15822 GCGGCAACTGAACTCATGTATGC 1604

RESULT 37
 AAA69168 standard; DNA; 56506 BP.
 XX ID AAA69168 (first entry)

Human foetal cDNA, SEQ ID NO: 742.

KW Human; foetal protein; cytotropic; immunosuppressive; immunostimulant; nootropic; neuroprotective; thrombolytic; osteoplastic; antiinflammatory; gene therapy; antisense therapy; cancer; immune disorder;
 growth disorder; osteoporosis; thrombolytic disorder;
 nervous system disorder; inflammation; expressed sequence tag; EST; 86.

XX OS Homo sapiens.
 XX WO200153339-A2.
 XX 02-AUG-2001.
 XX PD 25-JAN-2001; 2001WO-US002723.
 XX PR 25-JAN-2000; 2000US-00491404.
 XX PR 15-SEP-2000; 2000US-00663870.
 XX PR 06-NOV-2000; 2000US-00707351.
 XX PA (HYSEQ INC.)
 XX PI Young G, Ford JB, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;
 PI Liu C, Asundi V, Zhou P, Werthman T;
 XX DR WPI; 2001-465571/50.
 XX P-PSDB; AAM06538.
 XX Novel fetal proteins useful for the treatment and diagnosis of diseases
 associated with dysfunction of the protein e.g. cancers, immune
 disorders, growth disorders, thrombolytic disorders, nervous system
 disorders and inflammation.
 XX PS Claim 1; Page 458; 71pp; English.
 XX The invention relates to novel foetal polypeptides encoded by
 polynucleotides comprising one of 477 sequences fully defined in the
 specification. The foetal polynucleotides and polypeptides are useful in
 the treatment and diagnosis of diseases such as cancers, immune
 disorders, growth disorders (e.g. osteoporosis), thrombolytic disorders,
 nervous system disorders and inflammation. The present sequence was
 assembled using an expressed sequence tag (EST) found to be expressed in
 human foetal tissue cDNA libraries as the seed
 Sequence 392 BP; 104 A; 93 C; 76 G; 119 T; 0 U; 0 Other;
 XX Score 18; DB 5; Length 392;
 XX Best Local Similarity 66.7%; Pred. No. 1.2e+02;
 XX Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 XX Qy 2 CGGCCGAAATGGAGTGTATGCTCT 27
 XX Db 113 CTGCTGAAATCAGTATTTCCT 138
 XX SQ Sequence 2000 BP; 491 A; 502 C; 447 G; 559 T; 0 U; 1 Other;
 XX Query Match 66.7%; Score 18;
 XX Best Local Similarity 80.8%; Pred. No. 1.5e+02;
 XX Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 XX Qy 1 GCGGCCGAAATGGAGTGTATGCTCT 26
 XX Db 1484 GCGCCGCAATTCCAGGTGTATC 1509
 XX RESULT 39
 ADJ41623 ID ADJ41623 standard; cDNA; 2000 BP.
 XX AC ADJ41623;
 XX DT 05-MAY-2004 (First entry)
 XX DE Plant cDNA #2623.
 XX KW Plant; gene; ss; transcription; plant genome augmentation; cereal;
 KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
 KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
 KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
 KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
 KW antifungal.
 XX Eukaryota.
 XX OS US2004016025-A1.
 XX PD 22-JAN-2004.
 XX PP 26-SEP-2002; 2002US-00260238.
 XX PR 26-SEP-2001; 2001US-0325277P.
 PR 26-SEP-2001; 2001US-0325448P.
 PR 04-APR-2002; 2002US-037060P.
 XX PA (BUDW) BUDWORTH P.
 PA (MOUGH) MOUGHAMER T.
 PA (BRIGG) BRIGGS S P.
 PA (COOP) COOPER B.
 PA (GLAZ) GLAZEBROOK J.
 PA (GOFF) GOFF S A.
 PA (KATA) KATAGIRI F.
 PA (KREP) KREBS J.
 PA (PROV) PROVART N.
 PA (RICK) RICKE D.
 PA (ZHUT) ZHUTU T.
 XX PI Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
 PI Goff SA, Katagiri F, Krebs J, Provar N, Ricke D, Zhu T;
 XX DR 2004-190374/18.
 XX PS Claim 1; SEQ ID NO 2623; 230pp; English.
 XX CC The invention relates to plant nucleotide sequences that direct seed-,
 CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
 CC or constitutive transcription of an operatively linked nucleic acid
 CC segment. The invention also relates to a method for augmenting a plant
 CC genome and a method of identifying a gene, where its expression is
 CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
 CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
 CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
 CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
 CC encode are useful for manipulating crop plants to alter or improve
 CC phenotypic characteristics, to produce large quantities of oil or
 CC proteins, to incur resistance to insecticides, viruses or fungi, and to
 CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
 CC have a high nutritional value with reduced apical dominance or dwarfism,
 CC early flowering or altered metabolic pathways. This sequence represents a
 CC plant nucleic acid of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification but was obtained in
 CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
 XX SQ Sequence 2000 BP; 491 A; 502 C; 447 G; 559 T; 0 U; 1 Other;
 XX Query Match 66.7%; Score 18;
 XX Best Local Similarity 80.8%; Pred. No. 1.5e+02;
 XX Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 XX Qy 1 GCGGCCGAAATGGAGTGTATGCTCT 26
 XX Db 1484 GCGCCGCAATTCCAGGTGTATC 1509
 XX RESULT 40
 ABL24354/C
 ID ABL24354 standard; DNA; 4093 BP.
 XX AC ABL24354;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 24535.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 XX OS Drosophila melanogaster.
 XX

PN WO200171042-A2.
XX 27-SEP-2001.
PD XX 23-MAR-2001; 2001WO-US009231.
PF XX 23-MAR-2000; 2000US-0191637P.
PR PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
PI XX Venter JC, Adams M, Li PWD, Myers EW;
XX DR 2001-656860/75.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.
PT XX
PS Claim 1, SEQ ID NO 24535; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL3051), expressed DNA sequences (ABL0180-ABL16175) and the encoded proteins (ABBS7737-ABB7207). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 4093 BP; 1185 A; 920 C; 888 G; 1100 T; 0 U; 0 Other;
SQ Query Match 66.7%; Score 18; DB 4; Length 4093;
Best Local Similarity 80.8%; Pred. No. 1.7e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 2 CGGGCGCAATTGAGTTAATGATCCT 27
Db 2476 CGGCTTCATTGAGTGGATAACCT 2451

Search completed: January 7, 2005, 11:40:47
Job time : 209.579 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 7, 2005, 10:58:39 ; Search time 48.7895 Seconds

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing First 45 summaries

Database : Issued_Patents_NA:*

1:	/cgcn2_6/_ptodata/1/ina/5A_COMB.seq;*
2:	/cgcn2_6/_ptodata/1/ina/5B_COMB.seq;*
3:	/cgcn2_6/_ptodata/1/ina/6A_COMB.seq;*
4:	/cgcn2_6/_ptodata/1/ina/6B_COMB.seq;*
5:	/cgcn2_6/_ptodata/1/ina/PCTUS_COMB.seq;*
6:	/cgcn2_6/_ptodata/1/ina/backfiles1.seq;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	19	70.4	660	3	US-08-812-946A-1	Sequence 1, Appli
2	19	70.4	1514	4	US-09-023-655-1316	Sequence 1316, Ap
3	18	66.7	1664976	4	US-09-916-421B-1	Sequence 1, Appli
4	18	66.7	1664976	4	US-09-592-570-1	Sequence 1, Appli
5	17.8	65.9	2040	4	US-09-107-532A-2735	Sequence 2735, Ap
6	17.6	65.2	40	4	US-09-702-498A-3	Sequence 3, Appli
c 7	17.6	65.2	16001	4	US-09-702-498A-4	Sequence 4, Appli
c 8	17.6	65.2	16001	4	US-09-600-319-3	Sequence 3, Appli
9	17.4	64.4	6176	4	US-09-916-594-224	Sequence 224, App
c 10	17.4	64.4	6290	4	US-09-913-039-106	Sequence 106, APP
c 11	17	63.0	130	4	US-09-536-977-13	Sequence 13, Appli
c 12	17	63.0	131	1	US-08-324-243-17	Sequence 17, Appli
c 13	17	63.0	131	1	US-08-532-390-17	Sequence 17, Appli
c 14	17	63.0	131	3	US-08-717-294-17	Sequence 17, Appli
c 15	17	63.0	131	5	CIT-US95-1151-17	Sequence 17, Appli
c 16	17	63.0	374	4	US-09-536-977-43	Sequence 43, Appli
c 17	17	63.0	603	4	US-09-134-000C-196	Sequence 196, Appli
c 18	17	63.0	1277	4	US-09-536-977-45	Sequence 45, Appli
c 19	17	63.0	1277	4	US-09-536-977-47	Sequence 47, Appli
c 20	17	63.0	1277	4	US-09-536-977-49	Sequence 49, Appli
c 21	17	63.0	1277	4	US-09-536-977-51	Sequence 51, Appli
c 22	17	63.0	1323	4	US-09-475-515-35	Sequence 35, Appli
c 23	17	63.0	1387	4	US-09-475-515-34	Sequence 34, Appli
c 24	17	63.0	1431	4	US-09-536-977-73	Sequence 73, Appli
c 25	17	63.0	1453	4	US-09-475-515-33	Sequence 33, Appli
c 26	17	63.0	1632	1	US-08-324-243-34	Sequence 34, Appli
c 27	17	63.0	1632	1	US-08-532-390-34	Sequence 34, Appli

RESULT 1
US-08-812-946A-1
; Sequence 1, Application US/08812946A
; Patent No. 6221637
; GENERAL INFORMATION:
; APPLICANT: Taneaki HIDA et al.
; TITLE OF INVENTION: XANTHENE DERIVATIVES, THEIR PRODUCTION AND
; CORRESPONDENCE ADDRESS:
; ADDRESSE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,946A
; FILING DATE: March 4, 1997
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REINSTATEMENT NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-812-946A-1
Query Match Score 70.4%; DB 3; Length 660;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 9 AATTGAGTTATGTATCTTCT 27


```

NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (3122993)..(3122993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312226)..(312226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature

```

```

; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1

Query Match Score 18; DB 4; Length 1664976;
Best Local Similarity 80.8%; Pred. No. 78;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GGGGCCAATGAACTATATCC 26
Db 609314 GCAGCAGCTATGATGATATGTATCC 609339

RESULT 4
US-09-692-570-1
Sequence 1, Application US/09692570
Patent No. 6797466
GENERAL INFORMATION:
APPLICANT: Built et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6797466
TITLE OF INVENTION: Jannaschii
FILE REFERENCE: PB275C1
CURRENT APPLICATION NUMBER: US/09/692,570
CURRENT FILING DATE: 2003-01-14
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
PRIOR APPLICATION NUMBER: US 08/916,421
PRIOR FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc feature
LOCATION: (28257)..(28258)
FEATURE:
NAME/KEY: misc feature
LOCATION: (28222)..(28222)
FEATURE:
NAME/KEY: misc feature
LOCATION: (8473)..(8473)
FEATURE:
NAME/KEY: misc feature
LOCATION: (8480)..(8480)
FEATURE:
NAME/KEY: misc feature
LOCATION: (84812)..(84812)
FEATURE:
NAME/KEY: misc feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g

```

```

FEATURE: NAME/KEY: misc_feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (98339)..(98339)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (140948)..(140948)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (161385)..(161385)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (319261)..(319261)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (627081)..(627081)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE: NAME/KEY: misc_feature

```

LOCATION: (1130881)..(1130881)
 OTHER INFORMATION: n equals a, t, c, or g
 FEATURE: misc_feature
 NAME/KEY: misc_feature
 LOCATION: (1130988)..
 OTHER INFORMATION: n equals a, t, c, or g
 FEATURE: misc_feature
 NAME/KEY: misc_feature
 LOCATION: (1313224)..
 OTHER INFORMATION: n equals a, t, c, or g
 FEATURE: misc_feature
 NAME/KEY: misc_feature
 LOCATION: (1349473)..
 OTHER INFORMATION: n equals a, t, c, or g
 Query Match Score 66.7%; Score 18; DB 4; Length 166497;
 Best Local Similarity 80.8%; Pred. No. 78;
 Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 GCGCCGCAATTGAGATTGTATCC 26
 Db 609314 GCAGCAGCTTATGATGTATCC 609339

RESULT 5
 US-09-107-532A-2735
 Sequence 2735, Application US/09107532A
 Patent No. 6583275
 GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: LYNN A DOUCETTE-STAMM AND DAVID BUSH
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02454
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD/ROM ISO9660
 COMPUTER SYSTEM: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE ASCII

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107-532A
 FILING DATE: 30-Jun-1998
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Arnielio, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 2735:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2040 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: Circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE: misc_feature

LOCATION: (B) LOCATION 1..2040
 SEQUENCE DESCRIPTION: SEQ ID NO: 2735
 US-09-107-532A-2735
 Query Match Score 65.9%; Score 17.8; DB 4; Length 2040;
 Best Local Similarity 90.5%; Pred. No. 26;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 7 GCAATTGAAAGTTATGGTATCC 27
 Db 1060 GCAATCGAAGTTATTATCC 1080

RESULT 6
 US-09-702-498A-3
 Sequence 3, Application US/09702498A
 Patent No. 6746860
 GENERAL INFORMATION:
 APPLICANT: Tokusumi, Tsuyoshi
 APPLICANT: Iida, Akihiro
 APPLICANT: Hasegawa, Mamoru
 APPLICANT: Naga, Yoshiyuki
 TITLE OF INVENTION: Paramyxovirus vectors used for transfer
 TITLE OF INVENTION: of foreign genes
 FILE REFERENCE: 50026/025001
 CURRENT APPLICATION NUMBER: US/09/702-498A
 CURRENT FILING DATE: 2000-10-31
 PRIOR APPLICATION NUMBER: JP 2000-152726
 PRIOR FILING DATE: 2000-05-18
 NUMBER OF SEQ ID NOS: 34
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 40
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Artificially synthesized sequence derived from the
 Sendai virus
 US-09-702-498A-3

Query Match Score 65.2%; Score 17.6; DB 4; Length 40;
 Best Local Similarity 83.3%; Pred. No. 14;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 GCGGCCGAATTGAAAGTTATCC 24
 Db 15 GCGGCCGAATTGCAAGTATCTAT 38

RESULT 7
 US-09-702-498A-4/c
 Sequence 4, Application US/09702498A
 Patent No. 6746860
 GENERAL INFORMATION:
 APPLICANT: Tokusumi, Tsuyoshi
 APPLICANT: Iida, Akihiro
 APPLICANT: Hasegawa, Mamoru
 APPLICANT: Naga, Yoshiyuki
 TITLE OF INVENTION: Paramyxovirus vectors used for transfer
 TITLE OF INVENTION: of foreign genes
 FILE REFERENCE: 50026/025001
 CURRENT APPLICATION NUMBER: US/09/702-498A
 CURRENT FILING DATE: 2000-10-31
 PRIOR APPLICATION NUMBER: JP 2000-152726
 PRIOR FILING DATE: 2000-05-18
 NUMBER OF SEQ ID NOS: 34
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 4
 LENGTH: 40
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Artificially synthesized sequence derived from the

; OTHER INFORMATION: Sendai virus
US-09-702-493A-4

Query Match Score 17.6; DB 4; Length 40;
Best Local Similarity 83.3%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GCGGCCGAATTGAAAGTTATGTATCT 24
Db 26 GCGGCCGAATTGAAAGATGGCAGATATCTAT 3

RESULT 8
US-09-600-319-3/c
; Sequence 3, Application US/09600319
; GENERAL INFORMATION:
; APPLICANT: Owens, Gary
; TITLE OF INVENTION: Identification of a Smooth Muscle Cell (SMC) Specific Smooth Musc
; TITLE OF INVENTION: Myosin Heavy Chain (SM-MHC) Promoter/Enhancer
; FILE REFERENCE: 00241-03
; CURRENT APPLICATION NUMBER: US/09/600,319
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: PCT/US99/01038
; PRIOR FILING DATE: 1999-01-15
; PRIOR APPLICATION NUMBER: 60/071,300
; PRIOR FILING DATE: 1998-01-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 16011
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-600-319-3

Query Match Score 17.6; DB 4; Length 16011;
Best Local Similarity 83.3%; Pred. No. 50;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 GCGCCAATTGAAAGTTATGTATCC 27
Db 6333 GCTGGAATTGAAAGTTATGTACCAT 6310

RESULT 9
US-09-976-594-224
; Sequence 224, Application US/09976594
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: Perl Program
; SEQ ID NO 224
; LENGTH: 6176
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: misc feature
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 282397.77
US-09-976-594-224

Query Match Score 17.4; DB 4; Length 6176;
Best Local Similarity 77.8%; Pred. No. 51;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GCGGCCGAATTGAAAGTTATGTATCC 25
Db 52 GCGGCCGAATTGAAAGTTATGTATCC 28

RESULT 12
 US-08-324-243-17/c
 Patent No. 5786464 Application US/08324243
 GENERAL INFORMATION:
 APPLICANT: SEED, BRIAN
 TITLE OF INVENTION: OVEREXPRESSION OF MAMMALIAN AND VIRAL
 NUMBER OF SEQUENCES: 37
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30B
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/324,243
 FILING DATE: 19-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: CLARK, PAUL T
 REGISTRATION NUMBER: 30,162
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 SEQUENCE CHARACTERISTICS:
 LENGTH: 131 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-324-243-17

Query Match 63.0%; Score 17; DB 1; Length 131;
 Best Local Similarity 80.0%; Pred. No. 36;
 Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 GCGGCCGAATTGAAAGTTATGTATTC 25
 Db 131 GCGGCCGAATTGAAAGTTATGTATTC 107

RESULT 13
 US-08-324-243-17/c
 Sequence 17, Application US/08532390
 Patent No. 5795737
 GENERAL INFORMATION:
 APPLICANT: SEED, BRIAN
 APPLICANT: HAAS, JURGEN
 TITLE OF INVENTION: High Level Expression of Proteins
 NUMBER OF SEQUENCES: 40
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/717,294
 FILING DATE: 20-SEP-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILED DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Elbing, Karen L
 REGISTRATION NUMBER: 35,238
 REFERENCE DOCKET NUMBER: 00786/345001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-428-0200
 TELEFAX: 617-428-7045
 TELEX:
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 131 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Other

RESULT 14
 US-08-717-294-17/c
 Sequence 17, Application US/08717294
 ; Parent No. 6114143
 GENERAL INFORMATION:
 APPLICANT: SEED, BRIAN
 APPLICANT: HAAS, JURGEN
 TITLE OF INVENTION: HIGH LEVEL EXPRESSION OF
 PROTEINS
 NUMBER OF SEQUENCES: 110
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Clark & Elbing LLP
 STREET: 176 Federal Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/717,294
 FILING DATE: 20-SEP-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILED DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Elbing, Karen L
 REGISTRATION NUMBER: 35,238
 REFERENCE DOCKET NUMBER: 00786/345001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-428-0200
 TELEFAX: 617-428-7045
 TELEX:
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 131 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Other

US-08-717-294-17
 Query Match 63.0%; Score 17; DB 3; Length 131;
 Best Local Similarity 80.0%; Pred. No. 36;
 Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 GCGGCCGAATTGAAAGTTATGTATC 25
 Db 131 GCGCCGCGTGAAGCTTGATC 107

RESULT 15
 PC-US95-111511-17/C
 Sequence 17, Application PC/TUS9511511
 GENERAL INFORMATION:
 APPLICANT: SEED, BRIAN
 TITLE OF INVENTION: OVEREXPRESSION OF MAMMALIAN AND VIRAL
 TITLE OF INVENTION: PROTEINS
 NUMBER OF SEQUENCES: 37
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30B
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/11511
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: CLARK, PAUL T
 REGISTRATION NUMBER: 30-162
 REFERENCE/DOCKET NUMBER: 00786/2226001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 544-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 131 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 PCT-US95-111511-17

Query Match 63.0%; Score 17; DB 5; Length 131;
 Best Local Similarity 80.0%; Pred. No. 36;
 Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 GCGGCCGAATTGAAAGTTATGTATC 25
 Db 131 GCGCCGCGTGAAGCTTGATC 107

RESULT 16
 US-09-536-977-43/C
 Sequence 43, Application US/09536977
 GENERAL INFORMATION:
 APPLICANT: FONSGAARD, ANDERS
 TITLE OF INVENTION: METHOD FOR PRODUCING A NUCLEOTIDE SEQUENCE CONSTRUCT
 TITLE OF INVENTION: WITH OPTIMIZED CODONS FOR AN HIV GENETIC VACCINE BASED
 ON A PRIMARY, EARLY HIV ISOLATE AND SYNTHETIC ENVELOPE
 TITLE OF INVENTION: BX08 CONSTRUCTS
 FILE REFERENCE: 030307/0169
 CURRENT APPLICATION NUMBER: US/09/536,977
 CURRENT FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: 60/128,558
 PRIOR FILING DATE: 1999-04-09
 PRIOR APPLICATION NUMBER: DK PA1999 00427
 FILE REFERENCE: 030307/0169
 CURRENT APPLICATION NUMBER: US/09/536,977
 CURRENT FILING DATE: 2000-03-29
 NUMBER OF SEQ ID NOS: 118
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 45

US-08-717-294-17
 Query Match 63.0%; Score 17; DB 4; Length 374;
 Best Local Similarity 80.0%; Pred. No. 44;
 Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 GCGGCCGAATTGAAAGTTATGTATC 25
 Db 138 GCGCCGCGTGAAGCTTGATC 114

RESULT 17
 US-09-134-000C-196
 Sequence 196, Application US/09134000C
 GENERAL INFORMATION:
 APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 032796-032
 CURRENT APPLICATION NUMBER: US/09/134,000C
 CURRENT FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/055,778
 PRIOR FILING DATE: 1997-08-15
 NUMBER OF SEQ ID NOS: 6812
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 196
 LENGTH: 603
 TYPE: DNA
 ORGANISM: Enterococcus faecalis
 US-09-134-000C-196

Query Match 63.0%; Score 17; DB 4; Length 603;
 Best Local Similarity 80.0%; Pred. No. 49;
 Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 GCGGCCGAATTGAAAGTTATGTATC 25
 Db 538 GCGCCGCGTGAAGCTTGATC 562

RESULT 18
 US-09-536-977-45/C
 Sequence 45, Application US/09536977
 GENERAL INFORMATION:
 APPLICANT: FONSGAARD, ANDERS
 TITLE OF INVENTION: METHOD FOR PRODUCING A NUCLEOTIDE SEQUENCE CONSTRUCT
 TITLE OF INVENTION: WITH OPTIMIZED CODONS FOR AN HIV GENETIC VACCINE BASED
 ON A PRIMARY, EARLY HIV ISOLATE AND SYNTHETIC ENVELOPE
 TITLE OF INVENTION: BX08 CONSTRUCTS
 FILE REFERENCE: 030307/0169
 CURRENT APPLICATION NUMBER: US/09/536,977
 CURRENT FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: 60/128,558
 PRIOR FILING DATE: 1999-04-09
 PRIOR APPLICATION NUMBER: DK PA1999 00427
 FILE REFERENCE: 030307/0169
 CURRENT APPLICATION NUMBER: US/09/536,977
 CURRENT FILING DATE: 2000-03-29
 NUMBER OF SEQ ID NOS: 118
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 45

```

; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus
; FEATURE: CDS
; LOCATION: (1) .. (1275)
US-09-536-977-45

Query Match      63.0%; Score 17; DB 4; Length 1277;
Best Local Similarity 80.0%; Pred. No. 58; Score 17; DB 4; Length 1277;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy      1 GCGGCCGAATTGAAAGTTATGTCATC 25
Db      1041 GCGGCCGAATTGAAAGTTATGTCATC 1017

RESULT 21
US-09-536-977-51/C
; Sequence 51, Application US/09536977
; Patent No. 6649409
; GENERAL INFORMATION:
; APPLICANT: FOMSGAARD, ANDERS
; TITLE OF INVENTION: METHOD FOR PRODUCING A NUCLEOTIDE SEQUENCE CONSTRUCT
; WITH OPTIMIZED CODONS FOR AN HIV GENETIC VACCINE BASED
; ON A PRIMARY, EARLY HIV ISOLATE AND SYNTHETIC ENVELOPE
; BX08 CONSTRUCTS
; FILE REFERENCE: 030307/0169
; CURRENT APPLICATION NUMBER: US/09/536,977
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/128,558
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: DK PA1999 00427
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 118
; SEQ ID NO: 47
; SOFTWARE: PatentIn Ver. 2.1
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus
; US-09-536-977-51

Query Match      63.0%; Score 17; DB 4; Length 1277;
Best Local Similarity 80.0%; Pred. No. 58; Score 17; DB 4; Length 1277;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy      1 GCGGCCGAATTGAAAGTTATGTCATC 25
Db      1041 GCGGCCGAATTGAAAGTTATGTCATC 1017

RESULT 22
US-09-475-515-35/C
; Sequence 35, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Christopher
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SPILBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 030307/0169
; CURRENT APPLICATION NUMBER: US/09/536,977
; PRIOR APPLICATION NUMBER: 60/128,558
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO: 35
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 1323
; TYPE: DNA
;
```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: SP120.modsF162.delV2
US-09-475-515-35

Query Match 63.0%; Score 17; DB 4; Length 1323;
Best Local Similarity 80.0%; Pred. No. 58; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 5; Location: (1)..(1428)
US-09-536-977-73

Query Match 63.0%; Score 17; DB 4; Length 1431;
Best Local Similarity 80.0%; Pred. No. 59; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 5; Location: (1)..(1428)
US-09-536-977-73

RESULT 23
US-09-475-515-34/C
Sequence 34, Application US/09475515A
Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 34
LENGTH: 1387
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
OTHER INFORMATION: SP120.modsF162.delV2
US-09-475-515-34

Query Match 63.0%; Score 17; DB 4; Length 1387;
Best Local Similarity 80.0%; Pred. No. 59; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 5; Location: (1)..(1428)
US-09-536-977-73/C
Sequence 73, Application US/09536977
Patent No. 6649409
GENERAL INFORMATION:
APPLICANT: FOMSGAARD, ANDERS
TITLE OF INVENTION: METHOD FOR PRODUCING A NUCLEOTIDE SEQUENCE CONSTRUCT
WITH OPTIMIZED CODONS FOR AN HIV GENETIC VACCINE BASED
ON A PRIMARY, EARLY HIV ISOLATE AND SYNTHETIC ENVELOPE
TITLE OF INVENTION: BX08 CONSTRUCTS
FILE REFERENCE: 030107/0169
CURRENT APPLICATION NUMBER: US/09/536,977
CURRENT FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/128,558
PRIOR FILING DATE: 1999-04-09
PRIOR APPLICATION NUMBER: DK PA1999 00427
NUMBER OF SEQ ID NOS: 118
SOFTWARE: PatentIn Ver. 2.1

RESULT 24
US-09-536-977-73/C
Sequence 73, Application US/09536977
Patent No. 6649409
GENERAL INFORMATION:
APPLICANT: FOMSGAARD, ANDERS
TITLE OF INVENTION: METHOD FOR PRODUCING A NUCLEOTIDE SEQUENCE CONSTRUCT
WITH OPTIMIZED CODONS FOR AN HIV GENETIC VACCINE BASED
ON A PRIMARY, EARLY HIV ISOLATE AND SYNTHETIC ENVELOPE
TITLE OF INVENTION: BX08 CONSTRUCTS
FILE REFERENCE: 030107/0169
CURRENT APPLICATION NUMBER: US/09/536,977
CURRENT FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/128,558
PRIOR FILING DATE: 1999-03-29
NUMBER OF SEQ ID NOS: 118
SOFTWARE: PatentIn Ver. 2.1

RESULT 25
US-09-475-515-33/C
Sequence 33, Application US/09475515A
Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 33
LENGTH: 1453
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
US-09-475-515-33

Query Match 63.0%; Score 17; DB 4; Length 1453;
Best Local Similarity 80.0%; Pred. No. 59; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 5; Location: (1)..(1428)
US-09-536-977-73/C
Sequence 34, Application US/08324243
Patent No. 5786464
GENERAL INFORMATION:
APPLICANT: SEED, BRIAN
TITLE OF INVENTION: OVEREXPRESSION OF MAMMALIAN AND VIRAL
PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30B
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/324,243
 FILING DATE: 19-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: CLARK, PAUL T
 REGISTRATION NUMBER: 30,162
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEX: (617) 542-8906
 TELEX: 200154
 ZIP: 02110
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1632 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-324-243-34

Query Match 63.0%; Score 17; DB 1; Length 1632;
 Best Local Similarity 80.0%; Pred. No. 61;
 Matches 20; Conservative 0; Mismatches 5;
 Indels 0; Gaps 0;
 Qy 1 GCGGCCGAATTGAAAGTTATGATC 25
 Db 1230 GCGCCGGTGTGAAAGCTGTGCATC 1206

RESULT 27
 US-08-390-34/c
 Sequence 34, Application US/08532390
 Patent No. 5795737
 GENERAL INFORMATION:
 APPLICANT: SEED, BRIAN
 APPLICANT: HAAS, JURGEN
 TITLE OF INVENTION: High Level Expression of Proteins
 NUMBER OF SEQUENCES: 40
 CORRESPONDENCE ADDRESS:
 ADDRESS: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30B
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/532,390
 FILING DATE: 22-SEP-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/324,243
 FILING DATE: 19-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: LECH, KAREN F.
 REGISTRATION NUMBER: 35,238
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEX: 200154
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1632 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-532-390-34
 Query Match 63.0%; Score 17; DB 1; Length 1632;
 Best Local Similarity 80.0%; Pred. No. 61;
 Matches 20; Conservative 0; Mismatches 5;
 Indels 0; Gaps 0;
 Qy 1 GCGGCCGAATTGAAAGTTATGATC 25
 Db 1230 GCGCCGGTGTGAAAGCTGTGCATC 1206

RESULT 28
 US-08-717-294-34/c
 Sequence 34, Application US/08717294
 Patent No. 6114148
 GENERAL INFORMATION:
 APPLICANT: SEED, BRIAN
 APPLICANT: HAAS, JURGEN
 TITLE OF INVENTION: HIGH LEVEL EXPRESSION OF PROTEINS
 NUMBER OF SEQUENCES: 110
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Clark & Elbing LLP
 STREET: 176 Federal Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ FOR Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/717,294
 FILING DATE: 20-SEP-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Elbing, Karen L
 REGISTRATION NUMBER: 35,238
 REFERENCE/DOCKET NUMBER: 00786/345001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-428-0200
 TELEFAX: 617-428-7045
 TELELEX:
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1632 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-717-294-34

Query Match 63.0%; Score 17; DB 3; Length 1632;
 Best Local Similarity 80.0%; Pred. No. 61;
 Matches 20; Conservative 0; Mismatches 5;
 Indels 0; Gaps 0;
 Qy 1 GCGGCCGAATTGAAAGTTATGATC 25
 Db 1230 GCGCCGGTGTGAAAGCTGTGCATC 1206

RESULT 29
 PCT-US5-11511-34/c
 Sequence 34, Application PC/TUS9511511
 GENERAL INFORMATION:
 APPLICANT: SEED, BRIAN
 TITLE OF INVENTION: OVEREXPRESSION OF MAMMALIAN AND VIRAL PROTEINS

NUMBER OF SEQUENCES: 37
 ADDRESS LINE: Fish & Richardson
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110-2804

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.3.0B
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/11511
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: CLARK, PAUL T.
 REGISTRATION NUMBER: 30,162
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEX: (201)54 542-8906
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1632 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 PCT/US95/11511-34

Query Match Score 17; DB 5; Length 1632;
 Best Local Similarity 80.0%; Pred. No. 61;
 Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GCGGCCGAATTGAAAGTTATGTATC 25
 Db 1230 GCGCCGCGCTTGAGCTGTGCATC 1206

RESULT 30
 US-09-475-515-41/C
 Sequence 41, Application US/09475515A
 Patent No. 6602705

GENERAL INFORMATION:
 APPLICANT: BARNETT, Susan
 APPLICANT: ZUR MEGDE, Jan
 APPLICANT: LIU, Hong
 APPLICANT: HARTOG, Karin
 APPLICANT: GREER, Catherine
 APPLICANT: SELBY, Mark
 APPLICANT: WALKER, Christopher
 TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
 FILE REFERENCE: 1621.002
 CURRENT APPLICATION NUMBER: US/09/475,515A
 CURRENT FILING DATE: 1999-12-30
 NUMBER OF SEQ ID NOS: 90
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 41
 LENGTH: 1836
 TYPE: DNA
 ORGANISM: Artificial Sequence
 OTHER INFORMATION: Description of Artificial Sequence:
 US-09-475-515-41

Query Match Score 17; DB 4; Length 1836;
 Best Local Similarity 80.0%; Pred. No. 63;
 Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GCGGCCGAATTGAAAGTTATGTATC 25
 Db 936 GCGCCGCGCTTGAGCTGTGCATC 912

RESULT 32
 US-09-475-515-47/C
 Sequence 47, Application US/09475515A
 Patent No. 6602705

GENERAL INFORMATION:
 APPLICANT: BARNETT, Susan
 APPLICANT: ZUR MEGDE, Jan
 APPLICANT: SRIVASTAVA, Indresh
 APPLICANT: LIAN, Ying
 APPLICANT: HARTOG, Karin
 APPLICANT: GREER, Catherine
 APPLICANT: SELBY, Mark
 APPLICANT: WALKER, Christopher
 APPLICANT: LIU, Hong
 APPLICANT: HARTOG, Karin
 APPLICANT: GREER, Catherine
 APPLICANT: SELBY, Mark
 APPLICANT: WALKER, Christopher
 TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
 FILE REFERENCE: 1621.002
 CURRENT APPLICATION NUMBER: US/09/475,515A
 CURRENT FILING DATE: 1999-12-30
 NUMBER OF SEQ ID NOS: 90
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 47
 LENGTH: 1836
 TYPE: DNA
 ORGANISM: Artificial Sequence
 OTHER INFORMATION: Description of Artificial Sequence:
 US-09-475-515-47

Query Match Score 17; DB 4; Length 1836;
 Best Local Similarity 80.0%; Pred. No. 63;

OTHER INFORMATION: gp140.mut8.modsF162.delV1/V2
US-09-475-515-47

Query Match 63.0%; Score 17; DB 4; Length 1836;
Best Local Similarity 80.0%; Pred. No. 63;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GCGGCCGAATTGAAAGTTATGATC 25
Db 936 GCGCCGCAAGTGAAGCTGTGCATC 912

RESULT 33
US-09-536-977-67/C
Sequence 67 Application US/09536977
; Patent No. 669409
; GENERAL INFORMATION:
; APPLICANT: FOMSGAARD, ANDERS
; TITLE OF INVENTION: METHOD FOR PRODUCING A NUCLEOTIDE SEQUENCE CONSTRUCT
; WITH OPTIMIZED CODONS FOR AN HIV GENETIC VACCINE BASED
; ON A PRIMARY, EARLY HIV ISOLATE AND SYNTHETIC ENVELOPE
; TITLE OF INVENTION:
; TITLE OF INVENTION: BX08 CONSTRUCTS
; FILE REFERENCE: 030307/0169
; CURRENT APPLICATION NUMBER: US/09/536,977
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/128,558
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: DK PA1999 00427
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 118
; SEQ ID NO: 67
; TYPE: DNA
; SOFTWARE: PatentIn Ver. 2.1
; LENGTH: 1918
; ORGANISM: Human immunodeficiency virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (1905)

Query Match 63.0%; Score 17; DB 4; Length 1918;
Best Local Similarity 80.0%; Pred. No. 63;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GCGGCCGAATTGAAAGTTATGATC 25
Db 1041 GCGCCGCAAGTGAAGCTGTGCATC 1017

RESULT 34
US-09-475-515-37/C
Sequence 37 Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 38
; LENGTH: 1944
; TYPE: DNA

RESULT 35
US-09-475-515-38/C
Sequence 38 Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 39
; LENGTH: 1944
; TYPE: DNA

RESULT 36
US-09-475-515-40/C
Sequence 40 Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 37
; LENGTH: 1944
; TYPE: DNA

NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: 9P140.mut.modsFI62.delV2
US-09-475-515-40

Query Match Score 17; DB 4; Length 1944;
Best Local Similarity 80.0%; Pred. No. 63;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 1044 GCCGCCGAGTTGAGCTGTGCATC 1020

RESULT 37
US-09-475-515-43/c
Sequence 43, Application US/09475515A
Patent No. 6602705

GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 43

LENGTH: 1944
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
OTHER INFORMATION: 9P140.mut7.modsFI62.delV2
US-09-475-515-43

Query Match Score 17; DB 4; Length 1944;
Best Local Similarity 80.0%; Pred. No. 63;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 1044 GCCGCCGAGTTGAGCTGTGCATC 1020

RESULT 38
US-09-475-515-46/c
Sequence 46, Application US/09475515A
Patent No. 6602705

GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher

TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 46

LENGTH: 2025
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
OTHER INFORMATION: 9P140.modsSS162

Query Match Score 17; DB 4; Length 2025;
Best Local Similarity 80.0%; Pred. No. 64;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 1125 GCCGCCGAGTTGAGCTGTGCATC 1101

RESULT 39
US-09-475-515-36/c
Sequence 36, Application US/09475515A
Patent No. 6602705

GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher

TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 36

LENGTH: 2025
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
OTHER INFORMATION: 9P140.modsSS162

Query Match Score 17; DB 4; Length 2025;
Best Local Similarity 80.0%; Pred. No. 64;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 1125 GCCGCCGAGTTGAGCTGTGCATC 1101

RESULT 40
US-09-475-515-39/c
Sequence 39, Application US/09475515A
Patent No. 6602705

GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher

APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
CURRENT APPLICATION NUMBER: US/09/475,515A
FILE REFERENCE: 1621.002
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 39
LENGTH: 2025
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
OTHER INFORMATION: gp140.mut.modsF162
us-09-475-515-39

Query Match 63.0%; Score 17; DB 4; Length 2025;
Best Local Similarity 80.0%; Pred. No. 64;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 GCGCCGAAATGAAAGTTATGTTAC 25
Db 1125 GCGGCCGAGTGAAGCTGTGCCTC 1101

Search completed: January 7, 2005, 11:00:35
Job time : 53.7895 secs

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: January 7, 2005, 11:00:44 ; Search time 1147.74 Seconds

Total number of hits satisfying chosen parameters: 134.979 Million cell updates/sec

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% ; Maximum Match 100%

Listing First 45 summaries

Database : Published Applications NA:/*

1: /cgn2_6/ptodata/2/pubnra/US07_PUBCOMB.seq;*

2: /cgn2_6/ptodata/2/pubnra/PCT_NEW_PUB.seq;*

3: /cgn2_6/ptodata/2/pubnra/US06_NEW_PUB.seq;*

4: /cgn2_6/ptodata/2/pubnra/US07_PUBCOMB.seq;*

5: /cgn2_6/ptodata/2/pubnra/US07_PUBCOMB.seq;*

6: /cgn2_6/ptodata/2/pubnra/PTCTUS_PUBCOMB.seq;*

7: /cgn2_6/ptodata/2/pubnra/US08_NEW_PUB.seq;*

8: /cgn2_6/ptodata/2/pubnra/US09_PUBCOMB.seq;*

9: /cgn2_6/ptodata/2/pubnra/US09A_PUBCOMB.seq;*

10: /cgn2_6/ptodata/2/pubnra/US09C_PUBCOMB.seq;*

11: /cgn2_6/ptodata/2/pubnra/US09C_PUBCOMB.seq;*

12: /cgn2_6/ptodata/2/pubnra/US09_NEW_PUB.seq;*

13: /cgn2_6/ptodata/2/pubnra/US10A_PUBCOMB.seq;*

14: /cgn2_6/ptodata/2/pubnra/US10B_PUBCOMB.seq;*

15: /cgn2_6/ptodata/2/pubnra/US10C_PUBCOMB.seq;*

16: /cgn2_6/ptodata/2/pubnra/US10C_PUBCOMB.seq;*

17: /cgn2_6/ptodata/2/pubnra/US10_E_PUBCOMB.seq;*

18: /cgn2_6/ptodata/2/pubnra/US10_NEW_PUB.seq;*

19: /cgn2_6/ptodata/2/pubnra/US11_NEW_PUB.seq;*

20: /cgn2_6/ptodata/2/pubnra/US60_NEW_PUB.seq;*

21: /cgn2_6/ptodata/2/pubnra/US60_PUBCOMB.seq;*

13 18.2 67.4 3693 15 US-10-097-111-11 Sequence 11, Appl1
 14 18.2 67.4 56506 15 US-10-097-111-10 Sequence 10, Appl1
 15 18 66.7 1112 17 US-10-437-963-36729 Sequence 96729, A
 16 18 66.7 2007 16 US-10-260-238-2623 Sequence 21821, Ap
 17 17.8 65.9 2178 16 US-10-282-122A-21821 Sequence 21821, A
 C 18 17.8 65.9 3624 17 US-10-697-828-23 Sequence 23, Appl1
 C 19 17.8 65.9 4121 17 US-10-697-828-19 Sequence 19, Appl1
 C 20 17.8 65.9 5084 17 US-10-697-828-16 Sequence 16, Appl1
 C 21 17.8 65.9 48445 13 US-10-087-192-25 Sequence 25, Appl1
 C 22 17.6 65.2 40 9 US-09-956-277-21 Sequence 21, Appl1
 C 23 17.6 65.2 40 9 US-09-956-277-22 Sequence 22, Appl1
 C 24 17.6 65.2 40 10 US-09-966-930-21 Sequence 22, Appl1
 C 25 17.6 65.2 40 10 US-09-966-930-22 Sequence 22, Appl1
 C 26 17.6 65.2 40 15 US-10-316-538-21 Sequence 22, Appl1
 C 27 17.6 65.2 40 15 US-10-316-538-22 Sequence 22, Appl1
 C 28 17.6 65.2 40 15 US-10-316-535-21 Sequence 21, Appl1
 C 29 17.6 65.2 40 15 US-10-316-535-22 Sequence 22, Appl1
 C 30 17.6 65.2 40 17 US-10-784-710-3 Sequence 3, Appl1
 C 31 17.6 65.2 40 17 US-10-784-710-4 Sequence 4, Appl1
 C 32 17.6 65.2 2006 17 US-10-437-963-33246 Sequence 33246, A
 C 33 17.6 65.2 2541 16 US-10-264-049-044 Sequence 404, Appl1
 C 34 17.6 65.2 16011 14 US-10-057-726-16 Sequence 16, Appl1
 C 35 17.4 64.4 41 10 US-09-769-736-186 Sequence 186, App
 C 36 17.4 64.4 154 11 US-09-987-899-5599 Sequence 5599, Ap
 C 37 17.4 64.4 285 11 US-09-987-899-5598 Sequence 5598, Ap
 C 38 17.4 64.4 292 16 US-10-242-515A-0195 Sequence 1095, Ap
 C 39 17.4 64.4 292 16 US-10-085-783A-1095 Sequence 1095, Ap
 C 40 17.4 64.4 514 15 US-10-029-396-1911 Sequence 1391, Ap
 C 41 17.4 64.4 870 16 US-10-282-122A-15342 Sequence 15342, A
 C 42 17.4 64.4 1029 16 US-10-425-114-4668 Sequence 4668, Ap
 C 43 17.4 64.4 1208 15 US-10-027-632-24609 Sequence 264609,
 C 44 17.4 64.4 1208 15 US-10-027-632-24609 Sequence 264609,
 C 45 17.4 64.4 1714 18 US-10-425-115-66262 Sequence 68262, A

ALIGNMENTS

RESULT 1
 US-08-940-544-1
 / Sequence 1, Application US/08940544B
 / Publication No. US20020018783A1
 / GENERAL INFORMATION:
 / APPLICANT: SADELAIN, MICHEL
 / APPLICANT: CHEUNG, NAI-KONG V.
 / APPLICANT: KRAUSE, ANNA
 / APPLICANT: GUO, HONG-PEN
 / TITLE OF INVENTION: FUSION PROTEINS OF A SINGLE CHAIN ANTIBODY AND CD28 AND
 / FILE REFERENCE: MSK-P-035-US
 / CURRENT APPLICATION NUMBER: US/08/940,544B
 / CURRENT FILING DATE: 1997-09-30
 / EARLIER APPLICATION NUMBER: PCT/US97/04427
 / EARLIER FILING DATE: 1997-03-20
 / NUMBER OF SEQ ID NOS: 4
 / SOFTWARE: PatentIn Ver. 2.0
 / SEQ ID NO 1
 / LENGTH: 27
 / TYPE: DNA
 / ORGANISM: HUMAN
 / FEATURE:
 / OTHER INFORMATION: Upstream primer for PCR amplification.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description

1	27	100.0	27	8 US-08-940-544-1	Sequence 1, Appl1
2	19	70.4	328	16 US-10-448-256-6	Sequence 6, Appl
3	19	70.4	663	11 US-09-937-722-258	Sequence 258, Appl
4	19	70.4	1514	9 US-09-738-546-1	Sequence 1, Appl
5	19	70.4	1514	11 US-09-836-544-7	Sequence 7, Appl
6	19	70.4	1514	15 US-10-207-655-98	Sequence 98, Appl
7	19	70.4	1514	16 US-10-643-1316	Sequence 1316, Ap
8	19	70.4	3803	15 US-10-076-934-1	Sequence 1, Appl
9	19	70.4	3804	15 US-10-133-238-1	Sequence 257, Appl
10	19	70.4	3805	10 US-09-835-297-3	Sequence 1, Appl
11	19	70.4	51365	11 US-09-957-722-256	Sequence 3, Appl
12	19	70.4			Sequence 256, App

RESULT 2
US-10-448-256-6
Sequence 6, Application US/10448256
GENERAL INFORMATION:
APPLICANT: Sadelain, Michel
APPLICANT: Brentjens, Renier
APPLICANT: Maher, John
TITLE OF INVENTION: Chimeric T Cell Receptors
FILE REFERENCE: MSK P-058
CURRENT APPLICATION NUMBER: US/10/448, 256
CURRENT FILING DATE: 2003-05-28
PRIOR APPLICATION NUMBER: 60/383, 872
PRIOR FILING DATE: 2002-05-28
NUMBER OF SEQ ID NOS: 23
SEQ ID NO 6
LENGTH: 328
TYPE: DNA
ORGANISM: human
US-10-448-256-6

Query Match 70.4%; Score 19; DB 16; Length 328;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATGAAAGTTATGATCCCT 27
Db 4 AATGAAAGTTATGATCCCT 22

RESULT 3
US-09-997-722-258
Sequence 258, Application US/0997722
GENERAL INFORMATION:
APPLICANT: Morris, David
APPLICANT: Engelhard, Eric
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: A-71171/RMS/DCP
CURRENT APPLICATION NUMBER: US/09/997, 722
CURRENT FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: US/09/747, 377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US/09/798, 586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 301
SOFTWARE: PatentIn version 3.1
SEQ ID NO 258
LENGTH: 663
TYPE: DNA
ORGANISM: Homo sapiens
US-09-997-722-258

Query Match 70.4%; Score 19; DB 11; Length 663;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATGAAAGTTATGATCCCT 27
Db 339 AATGAAAGTTATGATCCCT 357

RESULT 4
US-09-738-546-1
Sequence 1, Application US/09738546
GENERAL INFORMATION:
APPLICANT: ANSETTI, CLAUDIO
TITLE OF INVENTION: CD28-SPECIFIC ANTIBODY COMPOSITIONS FOR USE IN METHODS
FILE REFERENCE: PHCC:0075

RESULT 5
US-09-836-544-7
Sequence 7, Application US/09836544
Publication No. US20040072283A1
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Rapid Immunoselection Cloning Method
FILE REFERENCE: 11-8L
CURRENT APPLICATION NUMBER: US/09/836, 544
CURRENT FILING DATE: 2001-04-17
PRIOR APPLICATION NUMBER: US/09/836, 544
CURRENT FILING DATE: 1992-12-01
PRIOR APPLICATION NUMBER: US/07/983, 647
PRIOR FILING DATE: 1990-07-13
PRIOR APPLICATION NUMBER: US/07/498, 809
PRIOR FILING DATE: 1990-03-23
PRIOR APPLICATION NUMBER: US/07/379, 076
PRIOR FILING DATE: 1989-07-13
PRIOR APPLICATION NUMBER: US/07/160, 416
PRIOR FILING DATE: 1988-02-25
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 1514

Query Match 70.4%; Score 19; DB 11; Length 1514;
Best Local Similarity 100.0%; Pred. No. 1.e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 AATGAAAGTTATGATCCCT 27
Db 438 AATGAAAGTTATGATCCCT 456

RESULT 6
US-10-207-655-98
Sequence 98, Application US/10207655
Publication No. US20030118592A1
GENERAL INFORMATION:
APPLICANT: Ledbetter, Jeffrey A.
APPLICANT: Hayden-Ledbetter, Martha S.
TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
FILE REFERENCE: 390069_401CI
CURRENT APPLICATION NUMBER: US/10/207, 655
CURRENT FILING DATE: 2002-07-25
NUMBER OF SEQ ID NOS: 426

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 98
; LENGTH: 1514
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-207-655-98

Query Match 70.4%; Score 19; DB 15; Length 1514;
Best Local Similarity 100.0%; Pred. No. 1.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; MisMatch 0; Gap 0;

Qy 9 AATTGAACTTATGATCTT 27
Db 438 AATTGAACTTATGATCTT 456

RESULT 7
US-10-641-643-1316
Sequence 1.316, Application US/10641643
Publication No. US20040077003A1
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
Susan G. Stuart
Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
GENE EXPRESSION
NUMBER OF SEQUENCES: 15.08
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/641,643
FILING DATE: 14-Aug-2003
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1316:
SEQUENCE CHARACTERISTICS:
LENGTH: 1514 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 9338444
SEQUENCE DESCRIPTION: SEQ ID NO: 1316 :
US-10-641-643-1316

Query Match 70.4%; Score 19; DB 16; Length 1514;
Best Local Similarity 100.0%; Pred. No. 1.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; MisMatch 0; Gap 0;

Qy 9 AATTGAACTTATGATCTT 27
Db 438 AATTGAACTTATGATCTT 456

RESULT 8
US-10-076-934-1
Sequence 1, Application US/10076934
Publication No. US2003017022A1
GENERAL INFORMATION:
APPLICANT: O'Hara Jr., Richard
Nagelein, Ann Marie
TITLE OF INVENTION: AGENTS THAT SPECIFICALLY BLOCK CD28-MEDIATED SIGNALING AND USES THEREFOR
FILE REFERENCE: GRN-028
CURRENT APPLICATION NUMBER: US/10/076,934
CURRENT FILING DATE: 2002-02-19
PRIOR APPLICATION NUMBER: 60/269,756
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO: 1
LENGTH: 3803
TYPE: DNA
ORGANISM: Homo sapiens
US-10-076-934-1

Query Match 70.4%; Score 19; DB 15; Length 3803;
Best Local Similarity 100.0%; Pred. No. 1.4e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; MisMatch 0; Gap 0;

Qy 9 AATTGAACTTATGATCTT 27
Db 560 AATTGAACTTATGATCTT 578

RESULT 9
US-09-997-722-257
Sequence 257, Application US/09997722
Publication No. US20040072154A1
GENERAL INFORMATION:
APPLICANT: Morris, David
Engehard, Eric
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: A-71171/RMS.DCF
CURRENT APPLICATION NUMBER: US/09/997,722
CURRENT FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 301
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 257
LENGTH: 3804
TYPE: DNA
ORGANISM: Homo sapiens
US-09-997-722-257

Query Match 70.4%; Score 19; DB 11; Length 3804;
Best Local Similarity 100.0%; Pred. No. 1.4e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; MisMatch 0; Gap 0;

Qy 9 AATTGAACTTATGATCTT 27
Db 561 AATTGAACTTATGATCTT 579

RESULT 10
US-10-143-238-1
Sequence 1, Application US/10143238
Publication No. US20030166502A1
GENERAL INFORMATION:
APPLICANT: Green, Jonathan M.
Shaw, Andrey S.
TITLE OF INVENTION: DIFFERENTIAL REGULATION OF T CELL SURVIVAL AND
PROLIFERATION
PRIOR FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: 60/418,238
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 1
LENGTH: 3804
TYPE: DNA
ORGANISM: Homo sapiens
US-10-143-238-1

```

```

; FILE REFERENCE: GNN-030
; CURRENT APPLICATION NUMBER: US/10/143,228
; CURRENT FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: US 60/290,097
; PRIOR FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 3804
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (223) ... (885)
US-10-143-238-1

Query Match 70.4%; Score 19; DB 15; Length 3804;
Best Local Similarity 100.0%; Pred. No. 1.4e+02; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0; Publication No. US20030086932A1
Qy 9 AATGAAAGTTATGTATCCT 27
Db 561 AATGAAAGTTATGTATCCT 579

RESULT 11
US-09-835-297-3
Sequence 3, Application US/09835297
Publication No. US20030086932A1
GENERAL INFORMATION:
APPLICANT: Bluestone, Jeffrey
APPLICANT: Collins, Mary
APPLICANT: Griffin, Mathew
APPLICANT: Kranz, David
TITLE OF INVENTION: SURFACE-BOUND ANTIGEN BINDING PORTIONS OF ANTIBODIES
TITLE OF INVENTION: THAT BIND TO CTLA4 AND USES THEREFOR
FILE REFERENCE: GNN-014CP
CURRENT APPLICATION NUMBER: US/09/835,297
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/196,851
PRIOR FILING DATE: 2000-04-12
NUMBER OF SEQ ID NOS: 4
SEQ ID NO 3
LENGTH: 3806
TYPE: DNA
ORGANISM: Homo sapiens
US-09-835-297-3

Query Match 70.4%; Score 19; DB 10; Length 3806;
Best Local Similarity 100.0%; Pred. No. 1.4e+02; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0; Publication No. US20040072154A1
Qy 9 AATGAAAGTTATGTATCCT 27
Db 561 AATGAAAGTTATGTATCCT 579

RESULT 12
US-09-987-722-256
Sequence 256, Application US/09997722
Publication No. US20040072154A1
GENERAL INFORMATION:
APPLICANT: Morris, David
APPLICANT: Engelhard, Eric
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: A-71171RMS/DCP
CURRENT APPLICATION NUMBER: US/09/997,722
CURRENT FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02

Query Match 70.4%; Score 19; DB 11; Length 51365;
Best Local Similarity 100.0%; Pred. No. 2.1e+02; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0; Publication No. US2003013877A1
Qy 9 AATGAAAGTTATGTATCCT 27
Db 30447 AATGAAAGTTATGTATCCT 30465

RESULT 13
US-10-097-111-11
Sequence 11, Application US/10097111
Publication No. US2003013877A1
GENERAL INFORMATION:
APPLICANT: PELLETIER, JERRY
APPLICANT: GROS, PHILIPPE
APPLICANT: DUBOW, MICHAEL
TITLE OF INVENTION: DNA SEQUENCES FROM S. PNEUMONIAE BACTERIOPHAGE EP1 THAT ENCODE ANTI-MICROBIAL POLYPEPTIDES
TITL OF INVENTION: DNA SEQUENCES FROM S. PNEUMONIAE BACTERIOPHAGE EP1 THAT ENCODE ANTI-MICROBIAL POLYPEPTIDES
FILE REFERENCE: 073406-06033
CURRENT APPLICATION NUMBER: US/10/097,111
CURRENT FILING DATE: 2002-07-24
PRIOR APPLICATION NUMBER: 09/676,412
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/157,218
PRIOR FILING DATE: 1999-09-30
NUMBER OF SEQ ID NOS: 552
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 3693
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
US-10-097-111-11

Query Match 67.4%; Score 18.2; DB 15; Length 3693;
Best Local Similarity 87.0%; Pred. No. 3.2e+02; Indels 3; Gaps 0;
Matches 20; Conservative 0; Mismatches 3; Publication No. US2003013877A1
Qy 4 GCGCAATTGAAGTTATGTATCCT 26
Db 1582 GCGCAACTGAAGTCAATGTATGC 1604

RESULT 14
US-10-097-111-10
Sequence 10, Application US/10097111
Publication No. US2003013877A1
GENERAL INFORMATION:
APPLICANT: PELLETIER, JERRY
APPLICANT: GROS, PHILIPPE
APPLICANT: DUBOW, MICHAEL
TITLE OF INVENTION: DNA SEQUENCES FROM S. PNEUMONIAE BACTERIOPHAGE EP1 THAT ENCODE ANTI-MICROBIAL POLYPEPTIDES
FILE REFERENCE: 073406-06033
CURRENT APPLICATION NUMBER: US/10/097,111
CURRENT FILING DATE: 2002-07-24
PRIOR APPLICATION NUMBER: 09/676,412
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/157,218
PRIOR FILING DATE: 1999-09-30
NUMBER OF SEQ ID NOS: 552
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 56506
TYPE: DNA

```

; ORGANISM: *Streptococcus pneumoniae*
 US-10-997-111-10
 Query Match Score 18.2; DB 15; Length 56506;
 Best Local Similarity 87.0%; Pred. No. 4.9e+02;
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 SEQ ID NO: 2623 LENGTH: 2000 TYPE: DNA
 FEATURE:
 NAME/KEY: N_region
 LOCATION: (1674)..(1674)
 OTHER INFORMATION: n = any nucleotide
 US-10-260-238-2623

RESULT 15
 US-10-437-963-96729 Application US/10437963
 Sequence 96729, Application US/10437963
 GENERAL INFORMATION:
 APPLICANT: La Rosa, Thomas J.
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 APPLICANT: Wu, Wei
 APPLICANT: Boukharov, Andrey A.
 APPLICANT: Barbazuk, Brad
 APPLICANT: Li, Ping
 TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-1 (53221)B
 CURRENT APPLICATION NUMBER: US/10/437,963
 CURRENT FILING DATE: 2003-05-14
 NUMBER OF SEQ ID NOS: 204966
 SEQ ID NO: 96729 LENGTH: 1112
 TYPE: DNA
 ORGANISM: *Oryza sativa*
 FEATURE:
 NAME/KEY: unsure
 LOCATION: (1)..(1112)
 OTHER INFORMATION: unsure at all n locations
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT4530_94799C.1
 US-10-437-963-96729

Query Match Score 18; DB 17; Length 1112;
 Best Local Similarity 80.8%; Pred. No. 3.3e+02;
 Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 SEQ ID NO: 752 LENGTH: 2000 TYPE: DNA
 FEATURE:
 NAME/KEY: unsure
 LOCATION: (1)..(1112)
 OTHER INFORMATION: unsure at all n locations
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT4530_94799C.1
 US-10-260-238-2623

RESULT 16
 US-10-260-238-2623
 Sequence 2623, Application US/10260238
 Publication No. US20040016025A1
 GENERAL INFORMATION:
 APPLICANT: Budworth, Paul R.
 APPLICANT: Moughamer, Todd G.
 APPLICANT: Briggs, Steven P.
 APPLICANT: Cooper, Bret
 APPLICANT: Glazebrook, Jane
 APPLICANT: Goff, Stephen A.
 APPLICANT: Katagiri, Fumiaki
 APPLICANT: Kreps, Joel
 APPLICANT: Provart, Nicholas
 APPLICANT: Ricke, Darrell
 APPLICANT: Zhu, Tong
 TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
 FILE REFERENCE: 60111-NP
 CURRENT APPLICATION NUMBER: US/10/260,238
 CURRENT FILING DATE: 2002-09-26
 PRIOR APPLICATION NUMBER: US 6/0/325,448

PRIOR FILING DATE: 2001-09-26
 PRIOR APPLICATION NUMBER: US 60/325,277
 PRIOR FILING DATE: 2001-09-26
 PRIOR APPLICATION NUMBER: US 60/370,620
 PRIOR FILING DATE: 2002-04-04
 NUMBER OF SEQ ID NOS: 6077
 SEQ ID NO: 2623 LENGTH: 2000 TYPE: DNA
 ORGANISM: *Oryza sativa*
 FEATURE:
 NAME/KEY: N_region
 LOCATION: (1674)..(1674)
 OTHER INFORMATION: n = any nucleotide
 US-10-260-238-2623

Query Match Score 18; DB 16; Length 2000;
 Best Local Similarity 80.8%; Pred. No. 3.6e+02;
 Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 SEQ ID NO: 1484 LENGTH: 2000 TYPE: DNA
 FEATURE:
 NAME/KEY: Xu_H
 LOCATION: (1)..(1484)
 OTHER INFORMATION: See File Wrapper or PALM.
 US-10-282-122A-21821
 Sequence 21821, Application US/1028212A
 Publication No. US20040029129A1
 GENERAL INFORMATION:
 APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Hasebeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu_H
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA.034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 21821 LENGTH: 2178 TYPE: DNA
 ORGANISM: *Enterococcus faecium*

US-10-282-122A-21621
 Query Match 65.9%; Score 17.8; DB 16; Length 2178;
 Best Local Similarity 90.5%; Pred. No. 4.5e+02; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 2;
 Db 2788 GCAAGTGAAGTTGATCCT 2768

RESULT 18
 US-10-697-828-23/C
 Sequence 23, Application US/10697828
 Publication No. US20040185546A1
 GENERAL INFORMATION:
 APPLICANT: Rosen, Steven
 APPLICANT: Lee, Jin Kyu
 APPLICANT: Hemmerich, Stefan
 TITLE OF INVENTION: Novel Glycosyl Sulfotransferases GST-4-alpha, GST-4-beta, & GST-6
 FILE REFERENCE: UCAL-138DIV
 CURRENT APPLICATION NUMBER: US/10/697,828
 CURRENT FILING DATE: 2003-10-29
 PRIOR APPLICATION NUMBER: 09/593,828
 PRIOR FILING DATE: 2000-06-13
 PRIOR APPLICATION NUMBER: 60/144,694
 PRIOR FILING DATE: 1999-07-20
 NUMBER OF SEQ ID NOS: 23
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 16
 LENGTH: 5084
 TYPE: DNA
 ORGANISM: mouse
 US-10-697-828-16

Query Match 65.9%; Score 17.8; DB 17; Length 5084;
 Best Local Similarity 90.5%; Pred. No. 5.2e+02; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 2;
 Qy 7 GCAATTGAACTTAATGATCCT 27
 Db 3288 GCAAGTGAAGTTGATCCT 3268

RESULT 21
 US-10-087-192-25
 Sequence 25, Application US/10087192
 Publication No. US20020182586A1
 GENERAL INFORMATION:
 APPLICANT: Morris, David W.
 APPLICANT: Engelhardt, Eric K.
 TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 TREATMENT OF CANCER
 FILE REFERENCE: 523452000122
 CURRENT APPLICATION NUMBER: US/10/087,192
 CURRENT FILING DATE: 2002-03-01
 PRIOR APPLICATION NUMBER: US 09/747,377
 PRIOR FILING DATE: 2001-12-22
 PRIOR APPLICATION NUMBER: US 09/798,586
 PRIOR FILING DATE: 2001-03-02
 NUMBER OF SEQ ID NOS: 2059
 SOFTWARE: FastSEQ for Windows' Version 4.0
 SEQ ID NO 25
 LENGTH: 48445
 TYPE: DNA
 ORGANISM: Mus musculus
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1..(48445))
 OTHER INFORMATION: n = A,T,C or G
 US-10-087-192-25

Query Match 65.9%; Score 17.8; DB 13; Length 48445;
 Best Local Similarity 90.5%; Pred. No. 7.4e+02; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 2;
 Qy 4 GCGCAATTGAACTTATGAT 24
 Db 47072 GCCTAAATTGAACTTATGAT 47092

US-10-697-828-19/C
 Sequence 19, Application US/10697828
 Publication No. US20040185546A1
 GENERAL INFORMATION:
 APPLICANT: Rosen, Steven
 APPLICANT: Lee, Jin Kyu
 APPLICANT: Hemmerich, Stefan
 TITLE OF INVENTION: Novel Glycosyl Sulfotransferases GST-4-alpha, GST-4-beta, & GST-6
 FILE REFERENCE: UCAL-138DIV
 CURRENT APPLICATION NUMBER: US/10/697,828
 CURRENT FILING DATE: 2003-10-29
 PRIOR APPLICATION NUMBER: 09/593,828
 PRIOR FILING DATE: 2000-06-13
 PRIOR APPLICATION NUMBER: 60/144,694
 PRIOR FILING DATE: 1999-07-20
 NUMBER OF SEQ ID NOS: 23
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 19
 LENGTH: 4121
 TYPE: DNA
 ORGANISM: mouse
 US-10-697-828-19

Query Match 65.9%; Score 17.8; DB 17; Length 4121;
 Best Local Similarity 90.5%; Pred. No. 5.e+02; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 2;

RESULT 22
US-09-366-277-21
Sequence 21, Application US/09966277
Patent No. US20020169306A1
GENERAL INFORMATION:
APPLICANT: KITAZATO, Kai
APPLICANT: SHU, Tsugumine
APPLICANT: KUMA, Hidekazu
APPLICANT: UEDA, Yasuji
APPLICANT: ASAKAWA, Makoto
APPLICANT: HASEGAWA, Mamoru
APPLICANT: IIDA, Akihiro
APPLICANT: TOKITOU, Fumino
APPLICANT: HIRATA, Takahiro
APPLICANT: INOUYE, Makoto
TITLE OF INVENTION: ENVELOPE GENE-DEFICIENT PARAMYXOVIRUS
FILE REFERENCE: 50026/028001
CURRENT APPLICATION NUMBER: US/09/966,277
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: PCT/JP00/03195
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: JP 2001/283451
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: JP 11/200739
PRIOR FILING DATE: 1999-05-18
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 21
LENGTH: 40
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificially Synthesized Sequence
US-09-966-277-21

Query Match 65.2%; Score 17.6; DB 9; Length 40;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 20; Conservative 4; Indels 0; Gaps 0;

Qy 1 GCGGCCGAATTGAAAGTTATGTAT 24
Db 15 GCGGCCGAATGGCAGATCTAT 38

RESULT 23
US-09-366-277-22/c
Sequence 22, Application US/09966277
Patent No. US20020169306A1
GENERAL INFORMATION:
APPLICANT: KITAZATO, Kai
APPLICANT: SHU, Tsugumine
APPLICANT: KUMA, Hidekazu
APPLICANT: UEDA, Yasuji
APPLICANT: ASAKAWA, Makoto
APPLICANT: HASEGAWA, Mamoru
APPLICANT: IIDA, Akihiro
APPLICANT: TOKITOU, Fumino
APPLICANT: HIRATA, Takahiro
APPLICANT: INOUYE, Makoto
TITLE OF INVENTION: ENVELOPE GENE-DEFICIENT PARAMYXOVIRUS
FILE REFERENCE: 50026/028001
CURRENT APPLICATION NUMBER: US/09/966,277
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: PCT/JP00/03195
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: JP 2001/283451
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: JP 11/200740
PRIOR FILING DATE: 1999-05-18
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 21
LENGTH: 40
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificially Synthesized Sequence
US-09-966-277-21

Query Match 65.2%; Score 17.6; DB 10; Length 40;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 20; Conservative 4; Indels 0; Gaps 0;

Qy 1 GCGGCCGAATTGAAAGTTATGTAT 24
Db 15 GCGGCCGAATGGCAGATCTAT 38

RESULT 25
US-09-366-930-22/c
Sequence 22, Application US/09966277
Patent No. US20020169306A1
GENERAL INFORMATION:
APPLICANT: KITAZATO, Kai
APPLICANT: SHU, Tsugumine
APPLICANT: KUMA, Hidekazu
APPLICANT: UEDA, Yasuji
APPLICANT: ASAKAWA, Makoto
APPLICANT: HASEGAWA, Mamoru
APPLICANT: IIDA, Akihiro
APPLICANT: TOKITOU, Fumino
APPLICANT: HIRATA, Takahiro
APPLICANT: INOUYE, Makoto
TITLE OF INVENTION: ENVELOPE GENE-DEFICIENT PARAMYXOVIRUS
FILE REFERENCE: 50026/028001
CURRENT APPLICATION NUMBER: US/09/966,277
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: PCT/JP00/03195
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: JP 2001/283451
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: JP 11/200739

APPLICANT: HASEGAWA, Mamoru
 APPLICANT: IIDA, Akihiro
 APPLICANT: INOUE, Makoto
 APPLICANT: HIRATA, Takahiro
 APPLICANT: INOUE, Makoto
 TITLE OF INVENTION: PARAMYXOVIRUS-DERIVED RNP
 FILE REFERENCE: 50026/029001
 CURRENT APPLICATION NUMBER: US/09/966,930
 CURRENT FILING DATE: 2001-09-18
 PRIOR APPLICATION NUMBER: PCT/JP00/03194
 PRIOR FILING DATE: 2000-05-18
 PRIOR APPLICATION NUMBER: JP 2001/283451
 PRIOR FILING DATE: 2001-09-18
 PRIOR APPLICATION NUMBER: JP 11/200740
 PRIOR FILING DATE: 1999-05-18
 NUMBER OF SEQ ID NOS: 64
 SEQ ID NO: 22
 LENGTH: 40
 SOFTWARE: FastSEQ for Windows Version 4.0
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Artificially Synthesized Sequence
 US-09-966-930-22

Query Match 65.2%; Score 17.6; DB 15;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 20; Conservative 0; Mismatches 4;
 Indels 0; Gaps 0;

Qy 1 GCGGCCGAATTGAAAGTTATGTAT 24
 Db 26 GCGGCCGAATTGAAAGTTATCTAT 3

RESULT 27
 US-10-316-538-22/C

; Sequence 22, Application US/10316538
 ; Publication No. US20030166252A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KITAZATO, Kai
 ; APPLICANT: SHU, Tsugumine
 ; APPLICANT: KUMA, Hidekazu
 ; APPLICANT: UEDA, Yasuji
 ; APPLICANT: ASAKAWA, Makoto
 ; APPLICANT: HASEGAWA, Mamoru
 ; APPLICANT: IIDA, Akihiro
 ; APPLICANT: HIRATA, Takahiro
 ; APPLICANT: INOUE, Makoto
 ; APPLICANT: TOKUSUMI, Yumiko

TITLE OF INVENTION: PARAMYXOVIRUS-DERIVED RNP
 FILE REFERENCE: D3-102PCT-USC1
 CURRENT APPLICATION NUMBER: US/10/316,538
 CURRENT FILING DATE: 2002-12-10
 PRIOR APPLICATION NUMBER: PCT/JP02/09558
 PRIOR FILING DATE: 2002-09-18
 PRIOR APPLICATION NUMBER: US 09/966,930
 PRIOR FILING DATE: 2001-09-27
 PRIOR APPLICATION NUMBER: JP 2001-283451
 PRIOR FILING DATE: 2001-09-18
 PRIOR APPLICATION NUMBER: PCT/JP00/03194
 PRIOR FILING DATE: 2000-05-18
 PRIOR APPLICATION NUMBER: JP 1999-200740
 NUMBER OF SEQ ID NOS: 68
 SEQ ID NO: 22
 LENGTH: 40
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: artificially
 OTHER INFORMATION: synthesized sequence

US-10-316-538-22

Query Match 65.2%; Score 17.6; DB 15;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 20; Conservative 0; Mismatches 4;
 Indels 0; Gaps 0;

Qy 1 GCGGCCGAATTGAAAGTTATGTAT 24
 Db 26 GCGGCCGAATTGAAAGTTATCTAT 3

RESULT 28
 US-10-316-538-21

; Sequence 21, Application US/10316535
 ; Publication No. US20030170266A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KITAZATO, Kai
 ; APPLICANT: SHU, Tsugumine
 ; APPLICANT: KUMA, Hidekazu
 ; APPLICANT: UEDA, Yasuji
 ; APPLICANT: ASAKAWA, Makoto
 ; APPLICANT: HASEGAWA, Mamoru
 ; APPLICANT: IIDA, Akihiro
 ; APPLICANT: TOKITOU, Fumino
 ; APPLICANT: HIRATA, Takahiro
 ; APPLICANT: TOKUSUMI, Tsuyoshi

; APPLICANT: INOUE, Makoto

APPLICANT: INOUE, Makoto
 APPLICANT: TOKUSUMI, Yumiko
 TITLE OF INVENTION: ENVELOPE GENE-DEFICIENT PARAMYXOVIRUS VECTOR
 FILE REFERENCE: D3-103PCT-USC1
 CURRENT APPLICATION NUMBER: US/10/316,535
 CURRENT FILING DATE: 2002-12-10
 PRIOR APPLICATION NUMBER: PCT/JP02/09558
 PRIOR FILING DATE: 2002-09-18
 PRIOR APPLICATION NUMBER: US 09/966,277
 PRIOR APPLICATION NUMBER: JP 2001-283451
 PRIOR APPLICATION NUMBER: PCT/JP00/03195
 SEQ ID NO: 21
 LENGTH: 40
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE: OTHER INFORMATION: Description of Artificial Sequence: artificially synthesized sequence
 US-10-316-535-21
 Query Match 65.2%; Score 17.6%; DB 15; Length 40;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 GCGGCCGCAATGAAAGTTATGTAT 24
 Db 15 GCGGCCGCAATGCAAGATCTAT 38

RESULT 29
 US-10-316-535-22/C
 Sequence 22, Application US/10316535
 GENERAL INFORMATION:
 PUBLICACION NO. US20030170266A1
 APPLICANT: KITAZATO, Kai
 APPLICANT: SHU, Tsugumine
 APPLICANT: KUMA, Hidekazu
 APPLICANT: UEDA, Yasuji
 APPLICANT: ASAKAWA, Makoto
 APPLICANT: HASEGAWA, Mamoru
 APPLICANT: IIDA, Akihiro
 APPLICANT: TOKITO, Fumino
 APPLICANT: HIRATA, Toshihiro
 APPLICANT: TOKUSUMI, Tsuyoshi
 APPLICANT: INOUE, Makoto
 TITLE OF INVENTION: ENVELOPE GENE-DEFICIENT PARAMYXOVIRUS VECTOR
 FILE REFERENCE: D3-103PCT-USC1
 CURRENT APPLICATION NUMBER: US/10/316,535
 CURRENT FILING DATE: 2002-12-10
 PRIOR APPLICATION NUMBER: PCT/JP02/09558
 PRIOR FILING DATE: 2002-09-18
 PRIOR APPLICATION NUMBER: US 09/966,277
 PRIOR APPLICATION NUMBER: JP 2001-283451
 PRIOR APPLICATION NUMBER: PCT/JP00/03195
 PRIOR FILING DATE: 2000-05-18
 PRIOR APPLICATION NUMBER: JP 1999-200739
 NUMBER OF SEQ ID NOS: 68
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 22
 LENGTH: 40
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE: OTHER INFORMATION: Description of Artificial Sequence: artificially synthesized sequence
 US-10-316-535-21
 Query Match 65.2%; Score 17.6%; DB 15; Length 40;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 GCGGCCGCAATGAAAGTTATGTAT 24
 Db 15 GCGGCCGCAATGCAAGATCTAT 38

RESULT 31
 US-10-784-710-4/C
 Sequence 4, Application US/10784710
 Publication No. US2004013762A1
 GENERAL INFORMATION:
 APPLICANT: TOKUSUMI, Tsuyoshi
 APPLICANT: IIDA, Akihiro
 APPLICANT: HASEGAWA, Mamoru
 APPLICANT: NAGAI, Yoshiyuki
 TITLE OF INVENTION: Paramyxovirus vectors used for transfer of foreign genes
 FILE REFERENCE: 50026/025001
 CURRENT APPLICATION NUMBER: US/10/784,710
 CURRENT FILING DATE: 2004-02-23
 PRIOR APPLICATION NUMBER: US/09/702,498A
 PRIOR FILING DATE: 2000-10-31
 PRIOR APPLICATION NUMBER: JP 2000-152726
 PRIOR FILING DATE: 2000-05-18
 NUMBER OF SEQ ID NOS: 34
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 3
 LENGTH: 40
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Artificially synthesized sequence derived from the Sendai virus
 US-10-784-710-3
 Query Match 65.2%; Score 17.6%; DB 17; Length 40;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 GCGGCCGCAATGAAAGTTATGTAT 24
 Db 15 GCGGCCGCAATGCAAGATCTAT 38

RESULT 32
 US-10-784-710-4/C
 Sequence 4, Application US/10784710
 Publication No. US2004013762A1
 GENERAL INFORMATION:
 APPLICANT: TOKUSUMI, Tsuyoshi
 APPLICANT: IIDA, Akihiro
 APPLICANT: HASEGAWA, Mamoru
 APPLICANT: NAGAI, Yoshiyuki
 TITLE OF INVENTION: Paramyxovirus vectors used for transfer of foreign genes
 FILE REFERENCE: 50026/025001
 CURRENT APPLICATION NUMBER: US/10/784,710
 CURRENT FILING DATE: 2004-02-23
 PRIOR APPLICATION NUMBER: US/09/702,498A
 PRIOR FILING DATE: 2000-10-31
 PRIOR APPLICATION NUMBER: JP 2000-152726
 PRIOR FILING DATE: 2000-05-18
 NUMBER OF SEQ ID NOS: 34
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 4

```

; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificially synthesized sequence derived from the
; OTHER INFORMATION: Sendai virus
US-10-784-710-4

Query Match      65.2%; Score 17.6; DB 17; Length 40;
Best Local Similarity 83.3%; Pred. No. 2.8e+02; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy   1  GCGGCATTGAACTTGAAGTTATGTAT 24
Db   26 GGGCGCATGGCGATCTAT 3

RESULT 32
US-10-437-963-33246/c
; Sequence 33246, Application US/10437963
; Publication No. US2014012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 33246
; LENGTH: 2006
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)::(2006)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4510_37377C.1
US-10-437-963-33246

Query Match      65.2%; Score 17.6; DB 17; Length 2006;
Best Local Similarity 83.3%; Pred. No. 5.4e+02; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy   4  GCCGCAATTGAACTTGAAGTTATGTATCCT 27
Db   1871 GGGCGATTGAACTTGAACCT 1848

RESULT 33
US-10-264-049-404
; Sequence 404, Application US/10264049
; Publication No. US20140005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 60/185569
; PRIOR APPLICATION NUMBER: PCT/US01/185569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 186
; LENGTH: 41

```

TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Primer
 US-09-769-736-186

Query Match 64.4%; Score 17.4; DB 10; Length 41;
 Best Local Similarity 77.8%; Pred. No. 3.7e+02;
 Matches 21; Conservative 0; Mismatches 6;
 Indels 0; Gaps 0;

Qy 1 GCGCCGAAATTGAAGTGTATCTCT 27
 Db 3 GCGCCGAAATTGAAGTGTATCTCT 29

RESULT 36
 US-10-987-899-5599

Sequence 5599, Application US/09987899
 Publication No. US20040116682A1
 GENERAL INFORMATION:
 APPLICANT: Cheikh, Nordine
 APPLICANT: Miller, Phillip W.
 APPLICANT: O'Connell, Keith M.
 APPLICANT: Liu, Jingdong
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With the Carbon Assimilation Pathway
 FILE REFERENCE: 16517.258
 CURRENT APPLICATION NUMBER: US/09/987,899
 CURRENT FILING DATE: 2001-11-16
 PRIOR APPLICATION NUMBER: US 09/262,979
 PRIOR FILING DATE: 1999-03-04
 PRIOR APPLICATION NUMBER: US 60/076,712
 PRIOR FILING DATE: 1998-03-06
 NUMBER OF SEQ ID NOS: 7341
 SEQ ID NO 5599
 LENGTH: 154
 TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 OTHER INFORMATION: Clone ID: 701169557H1
 US-09-987-899-5599

Query Match 64.4%; Score 17.4; DB 11; Length 154;
 Best Local Similarity 77.8%; Pred. No. 4.5e+02;
 Matches 21; Conservative 0; Mismatches 6;
 Indels 0; Gaps 0;

Qy 1 GCGCCGAAATTGAAGTGTATCTCT 27
 Db 87 GAGGCCGAAATTCAAGCTTAATGCT 113

RESULT 37
 US-09-987-899-5598

Sequence 5598, Application US/09987899
 Publication No. US20040116682A1
 GENERAL INFORMATION:
 APPLICANT: Cheikh, Nordine
 APPLICANT: Miller, Phillip W.
 APPLICANT: O'Connell, Keith M.
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With the Carbon Assimilation Pathway
 FILE REFERENCE: 16517.258
 CURRENT APPLICATION NUMBER: US/09/987,899
 CURRENT FILING DATE: 2001-11-16
 PRIOR APPLICATION NUMBER: US 09/262,979
 PRIOR FILING DATE: 1999-03-04
 PRIOR APPLICATION NUMBER: US 60/076,712
 PRIOR FILING DATE: 1998-03-06
 NUMBER OF SEQ ID NOS: 7341
 SEQ ID NO 5598
 LENGTH: 285
 TYPE: DNA
 ORGANISM: Human

RESULT 38
 US-10-242-535A-1095/C

Sequence 1095, Application US/1024253A
 Publication No. US2004001365A1
 GENERAL INFORMATION:
 APPLICANT: Chondrodene Inc.
 APPLICANT: Liew, C. C.
 TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
 FILE REFERENCE: 4331/2005
 CURRENT APPLICATION NUMBER: US/10/242-535A
 CURRENT FILING DATE: 2002-09-12
 PRIOR APPLICATION NUMBER: US 10/085,783
 PRIOR FILING DATE: 2002-02-28
 PRIOR APPLICATION NUMBER: US 60/305,340
 PRIOR FILING DATE: 2001-07-13
 PRIOR APPLICATION NUMBER: US 60/275,017
 PRIOR FILING DATE: 2001-03-12
 PRIOR APPLICATION NUMBER: US 60/271,355
 PRIOR FILING DATE: 2001-02-28
 NUMBER OF SEQ ID NOS: 58994
 SEQ ID NO 1095
 LENGTH: 292
 TYPE: DNA
 ORGANISM: Human

RESULT 39
 US-10-242-535A-1095/C

Sequence 1095, Application US/1024253A
 Publication No. US2004001365A1
 GENERAL INFORMATION:
 APPLICANT: Chondrodene Inc.
 APPLICANT: Liew, C. C.
 TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
 FILE REFERENCE: 4331/2005
 CURRENT APPLICATION NUMBER: US/10/085,783A
 CURRENT FILING DATE: 2002-02-28
 PRIOR APPLICATION NUMBER: US 60/305,340
 PRIOR FILING DATE: 2001-07-13
 PRIOR APPLICATION NUMBER: US 60/275,017
 PRIOR FILING DATE: 2001-03-12
 PRIOR APPLICATION NUMBER: US 60/271,355
 PRIOR FILING DATE: 2001-02-28
 NUMBER OF SEQ ID NOS: 58994
 SEQ ID NO 1095
 LENGTH: 292
 TYPE: DNA
 ORGANISM: Human

US-10-085-781A-1095

Query Match 64.4%; Score 17.4; DB 16; Length 292;
 Best Local Similarity 77.8%; Pred. No. 5e-02; Mismatches 6; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 0;

Qy 1 GCGCCGCAATTGAAGTTATGTATCCT 27
 Db 285 GCTGttGACTTAAGTttATACCT 259

RESULT 40

US-10-029-386-1391

Sequence 1391, Application US/10029386

; Publication No. US20030194704A1

GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GENE EXPRESSION ANALYSIS TWO

FILE REFERENCE: AEONICA-X-2

CURRENT APPLICATION NUMBER: US/10/029,386

CURRENT FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOC: 31288

SOFTWARE: Annonax Sequence Listing Engine vers. 1.1

SEQ ID NO 1391

LENGTH: 514

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

; OTHER INFORMATION: MAP TO AL133515.2

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.52

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.72

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.78

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.65

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.59

; OTHER INFORMATION: NT HIT: AF156100.1, EVALUE 0.00e+00

; OTHER INFORMATION: SWISSPROT HIT: P34886, EVALUE 2.10e+00

; OTHER INFORMATION: EST_HUMAN HIT: AA227235.1, EVALUE 1.80e-01

US-10-029-386-1391

Query Match 64.4%; Score 17.4; DB 15; Length 514;
 Best Local Similarity 94.7%; Pred. No. 5.5e+02; Mismatches 1; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 1;

Qy 7 GCAATTGAAGTTATGTATC 25

Db 434 GCATTGAAGTTATATTC 452

Search completed: January 7, 2005, 14:32:48

Job time : 1149.74 secs

Run on : January 7, 2005, 10:58:39 ; Search time 1611.95 Seconds
 OM nucleic - nucleic search, using sw model
 Copyright (c) 1993 - 2005 Compugen Ltd.
 GenCore version 5.1.6
 (without alignments)
 610363 Million cell updates/sec

Title: US-09-786-502A-7
Perfect Score: 27
Sequence: 1 gcgccgcaatgtaaaggttatgtatcct 27

Scoring table: IDENTITY_NUC
Gapopen 10.0 , Gapext 1.0
Searched: 32822875 seqs, 1821985908 residues
Total number of hits satisfying chosen parameters:

Minimum DB rec length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%

ALIGNMENTS

FSCB-processing: Minimum match 9%
Maximum Match 100%
Listing first 45 summaries

Database :	EST:*
	1: gb_est1:*
	2: gb_est2:*
	3: gb_htc:*
	4: gb_est3:*
	5: gb_est4:*
	6: gb_est5:*
	7: gb_est6:*
	8: gb_gse1:*
	9: gb_gbs2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			DB	ID	Description
		Match	Length	DB			
c 1	19.6	72.6	866	2	BF213303	BF213303	601845014
c 2	19.2	71.1	852	5	BT034454	BQ734454	AGENCOURT
c 3	19	70.4	304	5	BT031941	BQ319411	MR0-CT045
c 4	19	70	446	4	BJ031520	BJ031520	BK031520
c 5	19	70	450	5	BX101583	BX101593	BK101593
c 6	19	70	510	2	BW949936	BST162006	CD691494
c 7	19	70	596	6	CD691494	EST8017	h
c 8	19	70	611	9	AY402311	Pan trogl	AY402331
c 9	19	70	654	5	BK55995	DKFZp_81F	AY402330
c 10	19	70	663	9	AY402330	Homo sapi	BX854672
c 11	19	70	710	5	BX854672	BX854672	BX854672
c 12	19	70	723	9	Mus muscu	AY954205	EST366275
c 13	19	70	730	2	AW954205	CKB00588	CKB00588
c 14	19	70	886	7	CK800588	CK800588	AGENCOURT
c 15	19	70	935	4	BT034026	BHSMEL1002	BHSMEL1002
c 16	19	70	1000	5	BX403908	BX403908	BX403908
c 17	18.6	68.9	521	6	CB822552	CB822552	LINBET87h
c 18	18.6	68.9	687	8	CB232047	BZ232047	h266G07..9
c 19	18.6	68.9	815	6	CB901075	CB901075	trico25xf
c 20	18.6	68.9	815	7	CF870869	CF870869	trico25xf
c 21	18.6	68.9	818	8	BH463620	BH463620	BH463620
c 22	18.6	68.9	831	5	BP512240	BP512240	BP512240
c 23	18.6	68.9	919	9	CG303571	CG303571	CG303571
c 24	18.6	68.9	919	6	CNC00271	CNC00271	CG16A0867V

ORIGIN Alto, CA."

Query Match Score 19 6%; DB 2; Length 866;
Best Local Similarity 84.6%; Pred. No. 1.5e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GCGCGCATTGAGTTATGTATCC 26
Db 803 GCGCCCACTTGGAGTTATGTATCC 778

RESULT 2
BQ734454 LOCUS BQ734454 852 bp mRNA linear EST 16-JUL-2002
DEFINITION IMAGE:5572858 5', mRNA sequence.
ACCESSION BQ734454
VERSION BQ734454.1
KEYWORDS EST
SOURCE Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.

REFERENCE
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordim, S., Costa, F.F., Goldman, G.H., Carvalho, M., Matsukura, A., Baia, G.S., Simpson, D.H., Brunstein, A., O'Hare, M.J., Soares, P.R., Brentani, R.R., Reis, I.F., de Souza, S.J. and Simpson, A.J.

COMMENT Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

PUBLISHED 20020663
MEDLINE 10737800
COMMENT Contact: Simpson A.J.G.
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.p?1=IR0&c2=MRO-CT0451-221100-022-b01&t3=2000-11-22&t4=1>)
Seq primer: puc 18 forward
High quality sequence stop: 6.
Location/Qualifiers 1..304
Source /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="Taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0451"
/note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from: ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the PUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES ORIGIN
Tissue Procurement: Dr. Igor David
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:
www-bio.llnl.gov/bbrrp/image/image.html
Plate: LLNL12318 row: p column: 11
High quality sequence stop: 610.
Location/Qualifiers 1..852
Source /organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="Taxon:83355"
/clone="IMAGE:5572858"
/dev_stage="embryo, stage 31-32"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHID XGC Emb4"
/note="Organ: whole embryo; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dN. Average insert size 2.1 kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."

RESULT 4
BJ031520/C
LOCUS BJ031520
DEFINITION BJ031520 NIBB Mochii normalized Xenopus neurula library Xenopus laevis cDNA clone XI00711 5', mRNA Sequence.
VERSION BJ031520.1
KEYWORDS EST,
Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anur; Mesobatrachia; Pipoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.

REFERENCE
AUTHORS Kitayama, A., Teraoka, C., Mochii, M., Ueno, N., Shin-i, T. and Kohara, Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)

RESULT 3
BQ319411
DEFINITION BQ319411-221100-022-b01 CTC0451
ACCESSION MRO-CT0451
VERSION BQ319411
KEYWORDS EST

Query Match Score 19 1%; DB 5; Length 852;
Best Local Similarity 87.5%; Pred. No. 2.4e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GGCGCAATTGAACCTTATGATTC 26
Db 606 GGCCCCAATGAACTTATGATTC 629

RESULT 4
BQ319411
DEFINITION BQ319411-221100-022-b01 CTC0451
ACCESSION MRO-CT0451
VERSION BQ319411
KEYWORDS EST

Query Match Score 19 2%; DB 5; Length 852;
Best Local Similarity 87.5%; Pred. No. 2.4e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GGCGCAATTGAACCTTATGATTC 26
Db 606 GGCCCCAATGAACTTATGATTC 629

COMMENT	FEATURES	SOURCE	URL	LOCATION/QUALIFIERS
Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6055 Fax: 81-559-81-6055 Email: tshini@genes.nig.ac.jp The information of this clone is available through the following URL. http://xenopus.nibb.ac.jp/ .				<pre> 1. .446 /organism="Xenopus laevis" /mol_type="mRNA" /db_xref="taxon:8355" /clone="XL007011" /tissue_type="whole embryo" /dev_stage="stage 15" /clone_lib="NTIBB Mochii normalized Xenopus neurula library" </pre>

/note="Vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAACTGGGGCGCTAAATTTCCTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match	70.4%	Score 19;	DB 5;	Length 450;	
Best Local Similarity	100.0%	Pred. No.	2.8e+02;		
Matches 19;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	

Qy 9 AATTGAAGTTATGATCATCCT 27
Db 418 AATTGAAGTTATGATCATCCT 436

RESULT 6

LOCUS	AW949936	510 bp	mRNA	linear	EST 01-JUN-2000
DEFINITION	EST162006 MAGE resequences,	MAGA Homo sapiens cDNA,		mRNA sequence.	
ACCESSION	AW949936	MAGA Homo sapiens cDNA,		mRNA sequence.	
VERSION	AW949936.1	MAGA Homo sapiens cDNA,		mRNA sequence.	
KEYWORDS	EST	MAGA Homo sapiens cDNA,		mRNA sequence.	
SOURCE	Homo sapiens (human)	MAGA Homo sapiens cDNA,		mRNA sequence.	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	MAGA Homo sapiens cDNA,		mRNA sequence.	
REFERENCE	1 (bases 1 to 510)	MAGA Homo sapiens cDNA,		mRNA sequence.	
AUTHORS	Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and Quackenbush,J.	MAGA Homo sapiens cDNA,		mRNA sequence.	
TITLE	Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray	MAGA Homo sapiens cDNA,		mRNA sequence.	
JOURNAL	Unpublished (2000)	MAGA Homo sapiens cDNA,		mRNA sequence.	
COMMENT	Contact: John Quackenbush	MAGA Homo sapiens cDNA,		mRNA sequence.	
	The Institute for Genomic Research	MAGA Homo sapiens cDNA,		mRNA sequence.	
	9712 Medical Center Dr., Rockville, MD 20850, USA	MAGA Homo sapiens cDNA,		mRNA sequence.	
	Tel: 301 838 3528	MAGA Homo sapiens cDNA,		mRNA sequence.	
	Fax: 301 838 0208	MAGA Homo sapiens cDNA,		mRNA sequence.	
	Email: johnq@tigr.org	MAGA Homo sapiens cDNA,		mRNA sequence.	
	Plate: 8	MAGA Homo sapiens cDNA,		mRNA sequence.	
SEQ PRIMER:	Reverse.	MAGA Homo sapiens cDNA,		mRNA sequence.	
FEATURES	Location/Qualifiers	MAGA Homo sapiens cDNA,		mRNA sequence.	
source	1..510	MAGA Homo sapiens cDNA,		mRNA sequence.	
	/organism="Homo sapiens"	MAGA Homo sapiens cDNA,		mRNA sequence.	
	/mol_type="mRNA"	MAGA Homo sapiens cDNA,		mRNA sequence.	
	/db_xref="taxon:9606"	MAGA Homo sapiens cDNA,		mRNA sequence.	
	/clone_id=IMAGE resequences, MAGA"	MAGA Homo sapiens cDNA,		mRNA sequence.	
/note="Vector: pBluescriptSK"		MAGA Homo sapiens cDNA,		mRNA sequence.	
Query Match	70.4%	Score 19;	DB 2;	Length 510;	
Best Local Similarity	100.0%	Pred. No.	2.8e+02;		
Matches 19;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
LOCUS	CD691494	596 bp	mRNA	linear	EST 25-JUN-2000
DEFINITION	EST8017 human nasopharynx	Homo sapiens cDNA,		mRNA sequence.	
ACCESSION	CD691494	Homo sapiens cDNA,		mRNA sequence.	
VERSION	CD691494.1	Homo sapiens cDNA,		mRNA sequence.	
KEYWORDS	EST.	Homo sapiens (human)		mRNA sequence.	

ORGANISM Homo sapiens /locus_tag="HCM1183"

REFERENCE AY402331 Query Match Score 19; DB 9; Length 611;

AUTHORS Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and Best Local Similarity 100.0%; Pred. No. 2.9e+02;

Zeng,Y.-X. Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

TITLE Unpublished (2003) Qy 9 AATTGAGTTATGTATCCCT 27

JOURNAL COMMENT Contact: Yixin Zeng Db 287 AATTGAGTTATGTATCCCT 305

RESULT 9

FEATURES source LOCUS BX955995 654 bp mRNA linear EST 01-MAR-2004

DEFINITION DKFZD781F1075_r1_781 (synonym: hlc4c) Homo sapiens cDNA clone

ACCESSION DKFZD781F1075_5', mRNA sequence.

VERSION BX955995

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE AUTHORS Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R., Meewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.

VERSION EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R., et al.)

TITLE Unpublished (2003)

JOURNAL Contact: MIPS

COMMENT MIPS

INGOLSTADTER LANDSTR.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.

No 51 sequence available.

This clone (DKFZp781F1075) is available at the RZPD in Berlin. Please contact the RZPD, Resourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

LOCATION/QUALIFIERS Location/Qualifiers

1. organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZp781F1075"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="781 (synonym: hlc4c)"

/note="vector: pSport1_SF1; Site_1: SfiI; Site_2: SfiIB; CDNA-collection"

ORIGIN

REFERENCE AY402331 Query Match Score 19; DB 6; Length 596;

AUTHORS Ay402331 Best Local Similarity 100.0%; Pred. No. 2.9e+02;

DEFINITION Pan troglodytes CD28 gene, VIRTUAL TRANSCRIPT, partial sequence, Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ACCESSION AY402331.1

VERSION GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE AUTHORS Clark,A.G., Gilanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tannenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

VERSION 1 to 611

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES source LOCUS AY402330 663 bp DNA transcript

DEFINITION Homo sapiens CD28 gene, VIRTUAL TRANSCRIPT, partial sequence, AY402330 genomic survey sequence.

ACCESSION AY402330

VERSION 1.611

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

<1..>611

/gene="CD28"

KEYWORDS	GSS.	source	Homo sapiens (human)					
SOURCE	Homo sapiens							
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
REFERENCE	1 (bases 1 to 663)							
AUTHORS	Clark,A.G., Glianowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Taranenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.							
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios							
JOURNAL	Science 302 (5652), 1960-1963 (2003)							
PUBLMED	14671302							
REFERENCE	2 (bases 1 to 663)							
AUTHORS	Clark,A.G., Glianowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Taranenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.							
TITLE	Direct Submission							
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA							
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.							
FEATURES	source							
	gene							
ORIGIN								
Query Match	70.4%	Score	19;	DB	9;	Length	663;	
Best Local Similarity	100.0%	Pred.	No.	2.9e+02;				
Matches	19;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps 0;
Qy	9 AATTGAACTTATATCCCT 27							
Db	339 AATTGAACTTATATCCCT 357							
RESULT 11								
BX854672/c	BX854672 mRNA	LOCUS	710 bp	mRNA	Linear	EST	11-DEC-2003	
DEFINITION	BX854672 Kirschner embryo St10/14 Xenopus laevis cDNA clone.							
ACCESSION	I-MAG998F228575 ; IMAGE:3516309 5'							
VERSION	BX854672							
KEYWORDS	EST.							
SOURCE	Xenopus laevis (African clawed frog)							
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Xenophylophora; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodidae; Xenopus; Xenopus; Xenopus.							
REFERENCE	1 (bases 1 to 710)							
AUTHORS	Schroth,A., Korn,B. and Landgrebe,J.							
TITLE	Xenopus laevis UniGene Set 1 (RZPDLIB No.988)							
JOURNAL	Unpublished (2003)							
COMMENT	Contact: Ina Rolfs							
	RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH							
	Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany							
	RZPD; IMA9998F228575.							
	RZPDLIB; I.M.A.G.E. CDNA Clone Collection (amp- resistant) (RZPDLIB No.988); http://www.rzpd.de/cgi-bin/products/showLib.pl.cgi?responses?libNo=988 RZPDLIB Xenopus laevis UniGene Set 1 (RZPDLIB No.988) http://www.rzpd.de/cgi-bin/products/showLib.pl.cgi?responses?libNo=988 Contact: Ina Rolfs							
	RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH							
	Heubnerweg 6, D-14059 Berlin, Germany							
	Tel: +49 30 32639 101.							
Fax: +49 30 32639 111								
www.rzpd.de								
This clone is available royalty-free from RZPD; contact RZPD (clonet@rzpd.de) for further information. Seq primer: SP6, 5' ATTAGCTGACTATAG 3'.								
Location/Qualifiers								
1. .710								
/organism="Xenopus laevis"								
/mol_type="mRNA"								
/db_xref="taxon:8355"								
/clone="I-MAG998F228575 ; IMAGE:3516309"								
/tissue_type="pooled embryos (stage 10-14)"								
/lab_host="DH10B (phage-resistant)"								
/clone_lab="Kirschner embryo St10 14"								
/note="Vector: pCS2+; Site 1: Not I; Site 2: Sal I; Size-selected for average insert size 1.2 kb. Library was constructed and donated by M. Kirchner (Harvard Medical School)."								
FEATURES	source							
	Query Match	70.4%	Score	19;	DB	5;	Length	710;
	Best Local Similarity	91.5%	Pred.	No.	2.9e+02;			
	Matches	22;	Conservative	0;	Mismatches	5;	Indels	0;
								Gaps 0;
ORIGIN								
Qy	1 GCGGCCGAATGAAAGTTATATCCCT 27							
Db	353 GCTGCAGAAATGAAAGTTATATCCCT 327							
RESULT 12								
AG281788	AG281788 LOCUS	DEFINITION	Mus musculus molossinus DNA, clone:MSMG01-053J22.TJ, genomic survey sequence.					
ACCESSION	AG281788	VERSION	AG281788.1	GI	47854665			
		KEYWORDS	MSS.					
		SOURCE	Mus musculus molossinus					
		ORGANISM	Mus musculus molossinus					
		REFERENCE	Mammalia; Eutheria; Rodentia; Muridae; Murinae; Mus.					
		AUTHORS	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.					
		TITLE	BAC end Sequences of Library MSMG01					
		JOURNAL	Unpublished					
		REFERENCE	1 (bases 1 to 723)					
		AUTHORS	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.					
		TITLE	Direct Submission					
		JOURNAL	Submitted (17-NOV-2003) Massahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.jp), URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)					
		COMMENT	Clones are derived from the mouse BAC library MSMG01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp).					
			Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN), Koyada, Tsukuba, 305-0074 Japan, phone: 81-298-36-9189, fax: 81-298-36-9199					
		PRIMERS	e-mail: abe@rtc.riken.jp					
		SEQUENCING	: TJ					
		LIBRARY						
		Vector	: PBAC3.6					
		R.Site 1	: EcoRI					
		R.Site 2	: EcoRI.					
		Location/Qualifiers						
		1. .723						
		/organism="Mus musculus molossinus"						
		/mol_type="genomic DNA"						
		/sub_species="molossinus"						
		/db_xref="taxon:57486"						
		/clone="MSMG01-053J22.TJ"						

/sex="male"											TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)
/tissue_type="mixture of kidney and spleen"											JOURNAL		Unpublished (1999)
/clone_id=MSMG01 Mouse Male BAC Library"											COMMENT		Contact: Daniela S. Gerhard, Ph.D.
ORIGIN											COMMENT		Office of Cancer Genomics / NIH
Query Match 70.4%; Score 19; DB 9; Length 723;											COMMENT		Bldg. 31 Rm10A07 Bethesda, MD 20892
Best Local Similarity 81.5%; Pred. No. 2.9e+02;											COMMENT		Email: cgabbs-2@mail.nih.gov
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;											COMMENT		Tissue Procurement: Igor B. David
Qy 1 CGGGCGCAATTGAACTTATATCTT 27											COMMENT		CDNA Library Preparation: Express Genomics
Db 637 GCTGCCATTGAACTATAACCCT 663											COMMENT		DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:											COMMENT		http://image.lnl.gov
RESULT 13											COMMENT		Plate: LLM15081 row: o column: 22
AW954205											COMMENT		High quality sequence stop: 721.
AK954205											COMMENT		Location/Qualifiers
DEFINITION 730 bp mRNA resequences, MAGE Homo sapiens cDNA, mRNA sequence.											FEATURES		1..886
ACCESSION AV954205											SOURCE		/organism="Xenopus laevis"
VERSION AV954205.1											SOURCE		/mol_type="mRNA"
KEYWORDS EST.											SOURCE		/db_xref="Taxon:8355"
SOURCE Homo sapiens (human)											SOURCE		/clone="IMAGE:7203696"
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.											SOURCE		/tissue_type="Pooled samples from 6 adult male testis"
REFERENCE 1 (bases 1 to 730)											SOURCE		/lab_host="DH10B Tona,"
AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Quackenbush, J., Holt, I. B., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and											SOURCE		/clone_1ib="NICHD_XGC_Te2N"
TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray											SOURCE		/note="Oligo-dT primer: 5'-pGACTTGTCAGTCGAGGCCGCC(T)25-3', and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >1kb resulted in an average insert size of 1.15 kb. This primary, microquantitated library is normalized to Cot5 (non-normalized primary library is NICHD_XGC_Te2) and was constructed by Express Genomics (Frederick, MD). Note: this is a NIH_MGC library"
JOURNAL Unpublished (2000)											ORIGIN		
COMMENT Contact: John Quackenbush											ORIGIN		
The Institute for Genomic Research											ORIGIN		
9712 Medical Center Dr., Rockville, MD 20850, USA											ORIGIN		
Tel: 301 838 3528											ORIGIN		
Fax: 301 838 0208											ORIGIN		
Email: johnq@tigr.org											ORIGIN		
Plate: 61											ORIGIN		
FEATURES source											ORIGIN		
FEATURES source											ORIGIN		
Query Match 70.4%; Score 19; DB 2; Length 730;											ORIGIN		
Best Local Similarity 100.0%; Pred. No. 2.9e+02;											ORIGIN		
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;											ORIGIN		
Qy 9 AATTGAAGTTATGTATCCCT 27											ORIGIN		
Db 343 AATTGAAGTTATGTATCCCT 361											ORIGIN		
RESULT 14											ORIGIN		
CK800588/C											ORIGIN		
LOCUS CK800588											ORIGIN		
DEFINITION AGENCOURT 18833449 NICHD_XGC_Te2N Xenopus laevis cDNA clone IMAGE:7203696 5', mRNA sequence.											ORIGIN		
ACCESSION CK800588											ORIGIN		
VERSION CK800588.1											ORIGIN		
KEYWORDS EST.											ORIGIN		
SOURCE Xenopus laevis (African clawed frog)											ORIGIN		
ORGANISM Xenopus laevis											ORIGIN		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Xenopoda; Batrichia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopus; Xenopus; Xenopus.											ORIGIN		
REFERENCE NIH-MGC http://mgc.ncbi.nlm.nih.gov/											ORIGIN		
AUTHORS											ORIGIN		
RESULT 15											ORIGIN		
BI950426/C											ORIGIN		
LOCUS BI950426020K12f											ORIGIN		
DEFINITION Hordeum vulgare spike EST library HVCDNA0012 (Fusarium infected) Hordeum vulgare subsp. vulgare CDNA clone											ORIGIN		
Accession BI950426											ORIGIN		
Version BI950426.1											ORIGIN		
Keywords EST.											ORIGIN		
Organism Hordeum vulgare subsp. vulgare											ORIGIN		
Eukaryota; Viridiplantae; Streptophytina; Embryophyta; Tracheophyta; Poales; Poaceae; Poaceae; Triticeae; Hordeum.											ORIGIN		
Reference 1 (bases 1 to 35)											ORIGIN		
Authors Wing, R., Muehlbauer, G.J., Close, T.J., Kleinhofs, A., Wise, R., Heinen, S., Begum, D., Prisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simeone, J., Fenton, R.D., Matzras, M., Choi, D.W., Oates, R., and Main, D.											ORIGIN		
Title Development of a genetically and physically anchored EST resource for barley genomics: Fusarium infected Morex spike CDNA library											ORIGIN		
Journal Unpublished (2001)											ORIGIN		
Comment Contact: Wing RA Clemson University Genomics Institute Clemson University											ORIGIN		
100 Jordan Hall, Clemson, SC 29634, USA											ORIGIN		
REFERENCE NIH-MGC http://mgc.ncbi.nlm.nih.gov/											ORIGIN		
AUTHORS											ORIGIN		

Pax: 864 656 4293			
Email: rwing@clmson.edu			
Total hq bases = 369			
High quality sequence: 913.			
Location/Qualifiers			
1..935			
/organism="Hordeum vulgare subsp. vulgare"			
/mol_type="mRNA"			
/cultivar="Morex"			
/sub_species="vulgare"			
/db_xref="taxon:112509"			
/clone="HVSME10020K12f"			
/tissue_type="Spike"			
/lab_host="TJC121"			
/Clone lib="Hordeum vulgare spike EST library HVDNA0012 (Fusarium infected)"			
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XbaI; Plants were grown at the University of Minnesota in the GU Muchlauer lab; spikes were harvested and snap frozen at 0, 1, 2, 3, 4, 5, 6, and 8 days after Fusarium graminearum inoculation (Heinen). In the TJ Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all eight RNA pools were combined, poly(A) RNA was purified from the mixture, one primary unamplified cDNA library was made, and 1 million pfu were in vitro excised to give pBluescript SK(-) cDNA phagemids (Choi, Fenton, Malatrasi). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see htpp://www.genome.clemson.edu/orders Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (http://wheat.pw.usda.gov/9ggpages/bgn/31/cover.html)"			
FEATURES			
source			
/organism="Hordeum vulgare subsp. vulgare"			
/mol_type="mRNA"			
/cultivar="Morex"			
/sub_species="vulgare"			
/db_xref="taxon:112509"			
/clone="HVSME10020K12f"			
/tissue_type="Spike"			
/lab_host="TJC121"			
/Clone lib="Hordeum vulgare spike EST library HVDNA0012 (Fusarium infected)"			
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XbaI; Plants were grown at the University of Minnesota in the GU Muchlauer lab; spikes were harvested and snap frozen at 0, 1, 2, 3, 4, 5, 6, and 8 days after Fusarium graminearum inoculation (Heinen). In the TJ Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all eight RNA pools were combined, poly(A) RNA was purified from the mixture, one primary unamplified cDNA library was made, and 1 million pfu were in vitro excised to give pBluescript SK(-) cDNA phagemids (Choi, Fenton, Malatrasi). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see htpp://www.genome.clemson.edu/orders Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (http://wheat.pw.usda.gov/9ggpages/bgn/31/cover.html)"			
ORIGIN			
Query Match Score 19; DB 4; Length 935;			
Best Local Similarity 81.5%; Pred. No. 3e+02; Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;			
RESULT 16 BX403908/c			
LOCUS BX403908 Homo sapiens 1000 bp mRNA linear EST 03-MAY-2004			
DEFINITION Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone			
ACCESSION CL0BB0052E06			
VERSION 5-PRIME, mRNA sequence.			
KEYWORDS EST.			
SOURCE Homo sapiens (human)			
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Buteraria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE Li, W.B., Gruber, C., Jesse, J. and Polayes, D.			
TITLE Unpublished (2001)			
JOURNAL On May 15, 2003 this sequence version replaced gi:30768436.			
COMMENT Contact: Genoscope - Centre National de Sequencage Genoscope - Centre National de Sequencage			
FEATURES			
source			
/organism="Lotus japonicus var. japonicus"			
/mol_type="mRNA"			
/cultivar="Gifu (B-129)"			
/db_xref="taxon:34303"			
/dev_stage="5 and 7 week-old plants"			
/clone_id="Lotus japonicus nodule library 5 and 7 week-old"			
/note="Organ: Nodule; Vector: pSPORT1; Site_1: SalI; Site_2: NotI; The library was prepared using mRNA extracted from nodules of 5 and 7 week-old Lotus plants. Nodules were induced by, and contained Mesorhizobium			

Db	132	GGACTCAATTGAAAGATGATCTT	156
FEATURES	source		
ORGANISM	organism="Hypocrella jechorina"		
KEYWORDS	/mol type="mRNA"		
KEYWORDS	/strain="QW6"		
KEYWORDS	/db xref="taxon:51453"		
KEYWORDS	/clone="tric025xt21"		
KEYWORDS	/dev_stage="mycelia"		
KEYWORDS	/clone_lib="T_reesei mycelial culture, Version 6 October 2003"		
KEYWORDS	/note="Vector: PREP3Y; Site_1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."		
ORIGIN			
Query Match	68.9%; Score 18.6; DB 7; Length 815;		
Best Local Similarity	84.0%; Pred. No. 4.7e+02;		
Matches	Conservative 0; Mismatches 4; Indels 0; Gaps 0;		
JOURNAL	COMMENT		
REFERENCE	AUTHORS		
ACCESSION	TITLE		
VERSION	JOURNAL		
SOURCE	COMMENT		
ORGANISM	ORGANISM		
KEYWORDS	/mol type="mRNA"		
KEYWORDS	/strain="T01000B13"		
KEYWORDS	/db xref="taxon:3712"		
KEYWORDS	/clone="BOHQCT78"		
KEYWORDS	/clone_lib="BOHQ"		
KEYWORDS	/note="Vector: phoSI; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOSt1 using BstXI linkers"		
FEATURES	source		
ORGANISM	organism="Brassica oleracea"		
KEYWORDS	/mol type="genomic DNA"		
KEYWORDS	/strain="T01000B13"		
KEYWORDS	/db xref="taxon:3712"		
KEYWORDS	/clone="BOHQCT78"		
KEYWORDS	/clone_lib="BOHQ"		
KEYWORDS	/note="Vector: phoSI; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOSt1 using BstXI linkers"		
ORIGIN			
Query Match	68.9%; Score 18.6; DB 8; Length 818;		
Best Local Similarity	84.0%; Pred. No. 4.7e+02;		
Matches	Conservative 0; Mismatches 4; Indels 0; Gaps 0;		
JOURNAL	COMMENT		
REFERENCE	AUTHORS		
ACCESSION	TITLE		
VERSION	JOURNAL		
SOURCE	COMMENT		
ORGANISM	ORGANISM		
KEYWORDS	/mol type="genomic DNA"		
KEYWORDS	/strain="T01000B13"		
KEYWORDS	/db xref="taxon:3712"		
KEYWORDS	/clone="BOHQCT78"		
KEYWORDS	/clone_lib="BOHQ"		
KEYWORDS	/note="Vector: phoSI; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOSt1 using BstXI linkers"		
ORIGIN			
Query Match	68.9%; Score 18.6; DB 8; Length 818;		
Best Local Similarity	84.0%; Pred. No. 4.7e+02;		
Matches	Conservative 0; Mismatches 4; Indels 0; Gaps 0;		
JOURNAL	COMMENT		
REFERENCE	AUTHORS		
ACCESSION	TITLE		
VERSION	JOURNAL		
SOURCE	COMMENT		
ORGANISM	ORGANISM		
KEYWORDS	/mol type="genomic DNA"		
KEYWORDS	/strain="T01000B13"		
KEYWORDS	/db xref="taxon:3712"		
KEYWORDS	/clone="BOHQCT78"		
KEYWORDS	/clone_lib="BOHQ"		
KEYWORDS	/note="Vector: phoSI; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOSt1 using BstXI linkers"		

Email: whitelaw@tigr.org
 Seq primer: TP
 Class: shared ends.

FEATURES	source	Location/Qualifiers
	1. organism="Zea mays"	/mol_type="genomic DNA"
	/strain="B73"	/db_xref="taxon:4577"
	/clone_lib="ZM_0.7-1.5KB"	/note="Vector: PACSK; Site 1: HinclII; 0.7-1.5 kb methylation filtered genomic DNA library"

ORIGIN

Query Match	Score 18 6;	DB 9;	Length 919;
Best Local Similarity 68.9%;	Pred. No. 4.8e+02;		
Matches 21; Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;

Qy 3 GCGCCAAATTGAAATTATGTTATCTT 27
 Db 37 GGCCCCATATAAAGTTATCTT 61

RESULT 24
 CN0307G6H/C
 LOCUS CNS07G6H 486 bp DNA linear GSS 02-OCT-2001
 DEFINITION Anopheles gambiae GSS SP6 end of clone 18E09 of library NotreDame1 from strain PEST of Anopheles gambiae (African malaria mosquito), genomic survey sequence.
 ACCESSION AL609371
 VERSION AL609371.1
 KEWORDS
 SOURCE Anopheles gambiae (African malaria mosquito)
 ORGANISM Anopheles gambiae
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
 1 (bases 1 to 486)
 REFERENCE Roth, C.W., Brey, P.T., Ke, Z. and Collins, F.H.
 AUTHORS
 TITLE Direct Submission
 JOURNAL Submitted (01-OCT-2001) Genoscope - Centre National de Sequencage : BP 191 91015 France
 - Web : www.genoscope.cns.fr

REFERENCE

AUTHORS	JOURNAL	COMMENT
1. Roth, C.W., Brey, P.T., Ke, Z. and Collins, F.H.	Submitted (01-OCT-2001) BBMI, Institut Pasteur, 25, rue du Dr. Roux, Paris 75015, France	This clone is from an A. gambiae BAC library provided by F.H. Collins and sequenced by Genoscope in collaboration with the Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.

FEATURES

source	Location/Qualifiers
1. organism="Anopheles gambiae"	/mol_type="genomic DNA"
/strain="PEST"	/db_xref="taxon:7165"
/clone="18E09"	/clone_lib="NotreDame1"
	/note="end : SP6"

ORIGIN

Query Match	Score 18 4;	DB 9;	Length 486;
Best Local Similarity 68.1%;	Pred. No. 5.5e+02;		
Matches 19; Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

Qy 6 CGCAATTGAAATTGTTATGTCCT 25
 Db 472 CGCAATTGAAATTGTCATC 453

RESULT 25
 AF338890 628 bp DNA linear GSS 29-SEP-2000

DEFINITION	LOCUS	VERSION	KEYWORDS
IM0070G08F Mouse 10kb plasmid UGCI1M0070G08_F, genomic survey sequence.	AZ338890	AZ338890.1	G1:10412612
			GSS.
			Mus musculus (house mouse)
			ORGANISM Mus musculus
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
			REFERENCE 1. (bases 1 to 628)
			AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Isilan,H., Longacre,S., Mahmoud,M., Meinen,E., Pedersen,T., Reilly,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
			Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
			Unpublished (2000)
			Contact: Robert B. Weiss
			University of Utah Genome Center
			Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
			Tel: 801 585 5606
			Fax: 801 585 7177
			Email: ddunn@genetics.utah.edu
			Insert Length: 100000 Std Error: 0.00
			Plate: 0010 row: G column: 08
			Seq. Primer: CCTGTAAAACGACGGCCAGT
			Class: plasmid ends
			High quality sequence stop: 628.
			Location/Qualifiers
			1. 628
			/organism="Mus musculus" /mol_type="genomic DNA" /strain="C57BL/6J" /db_xref="Taxon:10090" /clone="UGCI1M0070G08"
			/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-" /clone_lib="Mouse 10kb Plasmid UGCI1M library" /note="Vector: PWDazn; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnarec/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA Polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The ligated adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pBD42 (gi 4732114 gb AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
			ORIGIN
			Query Match Score 18.4; DB 8; Length 628;
			Best Local Similarity 95.0%; Pred. No. 5.7e+02;
			Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy	8 CAATTGAAATTGTTATGTCCT 27		
Db	128 CAATTGAAATTGTCATC 147		

Ligate in double stranded adapter containing BsgI and BamHI sites [5'-ggccgtgcacccggatcggaaaaaaag]
[5'-aattcttttttttcgatccccggatcgacgc]"

ORIGIN							
Query Match	67.4%	Score 18.2;	DB 5;	Length 768;			
Best Local Matches	Similarity 87.0%;	Pred. No. 7.3e+02;					
Db	20; Conservative	0; Mismatches 3;	Indels 0;	Gaps 0;			
Qy	4 GCCGCAATTGAAAGTTATGTATCC 26						
Db	196 GCTGCAATTGAAAGTTACGTATCC 174						
RESULT 29							
LOCUS	BU335623/c	BU335623	788 bp	mRNA	EST 28-NOV-2002		
DEFINITION	603496349F1. CSEOCHN65 Gallus gallus cDNA clone ChEST408a13 5', mRNA sequence.						
ACCESSION	EU335623						
VERSION	EU335623.1	GI: 25843624					
KEYWORDS	EST.						
SOURCE	Gallus gallus (chicken)						
ORGANISM	Gallus gallus						
Phasianinae; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.							
REFERENCE	1 (bases 1 to 825)						
AUTHORS	Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.						
TITLE	A Comprehensive Collection of Chicken cDNAs						
JOURNAL	Curr. Biol. 12 (22), 1965-1969 (2002)						
MEDLINE	2233534						
PUBMED	12445392						
COMMENT	Contact: Simon Hubbard Department of Biomolecular Sciences University of Manchester Institute of Science and Technology (UMIST) PO Box 88, Manchester, M60 1QD, UK Tel: 01612008930 Fax: 01612360409 Email: Simon.Hubbard@umist.ac.uk.						
FEATURES	source						
source	1. '825						
	/organism="Gallus gallus"						
	/mol type="mRNA"						
	/strain="White Leghorn, Hisex"						
	/db_xref="taxon:9031"						
	/clone="CSEOCHN75"						
	/note="Organ: trunk; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adaptors, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."						
ORIGIN							
Query Match	67.4%	Score 18.2;	DB 5;	Length 825;			
Best Local Matches	Similarity 87.0%;	Pred. No. 7.3e+02;					
Db	20; Conservative	0; Mismatches 3;	Indels 0;	Gaps 0;	Indels 3;	Mismatches 0;	Gaps 0;
Qy	4 GCGGCAATTGAAAGTTATGTATCC 26						
Db	199 GCTGCAATTGAAAGTTACGTATCC 177						
RESULT 31							
CNS02NST	838 bp	DNA linear	GSS 01-SEP-2000				
LOCUS	CNS02NST						
DEFINITION	Tetraodon nigroviridis genome survey sequence T7 end of clone 150J10 of library G from Tetraodon nigroviridis. Genomic survey						
ORIGIN							
Query Match	67.4%	Score 18.2;	DB 5;	Length 788;			
Best Local Matches	Similarity 87.0%;	Pred. No. 7.3e+02;					
Db	20; Conservative	0; Mismatches 3;	Indels 0;	Gaps 0;			

TITLE	Rogers,J., Birney,E. and Hayashizaki,Y. Analysis of the mouse transcriptome based on functional annotation	REFERENCE 1 (bases 1 to 508)
JOURNAL	Nature 420, 563-573 (2002)	AUTHORS Jantasuriyarat C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kuwana,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
MEDLINE	22354683	TITLE Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea
PUBLISHED	124666851	JOURNAL Unpublished (2003)
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) Tel: +81-45-503-9222 Fax: +81-45-503-9216 Email: genome-res@gsc.riken.jp; URL: http://genome.gsc.riken.jp/ Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotsiki,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numasaki,R., Oono,M., Sakai,K., Sakuzume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watanuki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission Computer Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome 12, 673-677 (2001) Normalization and subraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN integrated Sequence analysis (RISA) system-384-format sequencing pipeline with 384 capillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Kirk W. Beissel (Boys Town National Research Hospital 555 North 30th Street Omaha, NE 68131 USA) whose assistance we gratefully acknowledge. Please visit our web site (http://genome.gsc.riken.go.jp) for further details.	COMMENT between rice and Magnaporthe grisea Contact: Rod Wing Arizona Genomics Institute University of Arizona Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ Tel: 520 626 3967 Fax: 520 621 9288 Email: http://genome.arizona.edu PCR PRIMERS FORWARD: gta aaa cga cgg cca gtg BACKWARD: gga aac agc tat gac cat g Plate: 13 row: 1 column: 20 Seq primer: gta aaa cga cgg cca gtg. FEATURES Location/Qualifiers 1. 508 source /organism="Oryza sativa (japonica cultivar-group)" /mol_type="mRNA" /cnlCivar="Nipponbare" /db_xref="ITaxon:39947" /clone="OSJNBE13120" /tissue_type="leaf" /dev_stage="3 week" /lab_host="DH10B" /clone_lib="OSJNEE" /note="Vector: pBluescript II KS+; Site 1: EcoRI; Site 2: XbaI; 24 hrs after innoculation with Rice Blast (70-15) in xhoI; ORIGIN
FEATURES	Query Match Score 18; DB 6; Length 508; Best Local Similarity 80.8%; Pred. No. 8.7e+02; Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0; Db	Query Match Score 18; DB 6; Length 508; Best Local Similarity 80.8%; Pred. No. 8.7e+02; Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0; Db
LOCATIONS	1. 205 /organism="Mus musculus" /mol_type="mRNA" /strand="+5'BL/6J" /db_xref="ITaxon:10090" /clone="F930021E21" /tissue_type="inner ear" /dev_stage="adult" /clone_lib="RIKEN full-length enriched, adult inner ear" Location/Qualifiers 1. 205	RESULT 35 AQ654959 LOCUS Sheared DNA-21K22-TR Sheared DNA trypanosoma brucei genomic clone DEFINITION Sheared DNA-21K22, genomic survey sequence. ACCESSION AQ654959 VERSION AQ654959.1 KEYWORDS GSS. SOURCE ORGANISM trypanosoma brucei trypanosoma brucei trypanosomatidae; trypanosoma; kinetoplastida; trypanosomatidae; trypanosoma; trypanozoa; euglenozoa; kinetoplastida; trypanosomatidae; trypanosoma. Unpublished (1999) COMMENT Other GSSE: Sheared DNA-21K22-TF Contact: Najib M. El-Sayed Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: nalsayed@igr.org Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
FEATURES	Query Match Score 18; DB 5; Length 205; Best Local Similarity 80.8%; Pred. No. 7.9e+02; Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0; Db	RESULT 34 CB676984 LOCUS OSJNE13120.f OSJNE13120.5', mRNA sequence DEFINITION Clone OSJNE13120 5', mRNA sequence ACCESSION CB676984 EWORD EST SOURCE ORGANISM Eukaryota: Viridiplantae: Streptophytina: Embryophytina: Tracheophytina: Poales; Poaceae;
FEATURES	Query Match Score 18; DB 5; Length 205; Best Local Similarity 80.8%; Pred. No. 7.9e+02; Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0; Db	RESULT 34 CB676984 LOCUS OSJNE13120.f OSJNE13120.5', mRNA linear EST 09-APR-2003 DEFINITION Clone OSJNE13120 5', mRNA sequence ACCESSION CB676984 EWORD EST SOURCE ORGANISM Eukaryota: Viridiplantae: Streptophytina: Embryophytina: Tracheophytina: Poales; Poaceae;

DNA library constructed at TIGR. Clones will be available for distribution through ARCC. Sheared DNA end sequences search page:
<http://www.tigr.org/tdb/mdb/cbdb/>.

Seq primer: M13-Reverse

Class: shotgun

Location/Qualifiers

1..561

/organism="Trypanosoma brucei"

/mol_type="genomic DNA"

/strain="TREU927/4 GUTat 10.1"

/db_xref="taxon:5691"

/clone="Sheared DNA 21K22"

/note="Vector: PUC18; Site 1: Smal; Constructed at The

Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give tight size distribution (approx 2 kb).

The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C.

(Making small insert libraries for whole genome shotgun sequencing projects. In Genome sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).

ORIGIN

Query Match 66.7%; Score 18; DB 8; Length 561;
 Best Local Similarity 80.8%; Pred. No. 8.8e+02;
 Matches 21; Conservative 5; Indels 0; Gaps 0;

Qy 2 CGGGCGATTAAGTTGATGTATCCT 27
 Db 339 CGGCCGCAATTGATGTAGTCATGCTCT 364

RESULT 36

LOCUS BQ135922 562 bp mRNA linear EST 23-APR-2002

DEFINITION NF023D01EC1F1012 Elicited cell culture *Medicago truncatula* cDNA
 Clone NF023D01EC 5', mRNA sequence.

ACCESSION BQ135922

VERSION BQ135922.1

COMMENT GI:202272046

SOURCE *Medicago truncatula* (barrel medic)

Medicago truncatula

Viriditae; Streptophytina; Embryophytina; Trachophyta; Spermatophytina; Magnoliophytina; Eudicots; Rosids; eudicots; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaeae; Medicago.

1 (bases 1 to 562)

Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation - Center for *Medicago* Genomics Research

Published (2000)

Contact: Dixon RA

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7102

Fax: 580 221 7380

Email: radixon@noble.org

Insert Length: 562

Std Error: 0.00

Row: 023

Column: 01

Seq primer: TCACACGAAAGCTATGAC.

ORIGIN

/organism="Medicago truncatula"

/mol_type="mRNA"

/db_xref="taxon:3880"

/clones="NF023D01EC"

/tissue_type="Cell cultures derived from root tissues"

/dev_stage="Cell suspensions were subcultured every 14

days. Cells were induced six days after subculture"
 /clone lib="Elicited cell culture"
 /note="Vector: Lambda Zap; Cells were induced with yeast cell wall extracts equivalent to 50ug/ml glucose in the final concentration. Samples were taken at 0.5, 1, 12 and 24 hours after induction. Equal amounts of RNA from each time point were pooled and used for mRNA isolation."

ORIGIN

Query Match 66.7%; Score 18; DB 5; Length 562;
 Best Local Similarity 80.8%; Pred. No. 8.8e+02;
 Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 CGGGCGATTAAGTTGATGTATCCT 27

Db 103 CAGGCCAATTGAATTGATGATCCT 128

RESULT 37

AQ644499

LOCUS AQ644499 599 bp DNA linear RPC193-DpnII Trypanosoma brucei Genomic clone
 Definition RPC193-DpnII-28M12, Genomic survey sequence.
 Accession AQ644499

VERSION AQ644499.1

COMMENT GI:51121209

KEYWORDS GSS,

SOURCE Trypanosoma brucei

ORGANISM Trypanosoma brucei

TRYpanosoma

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

TRYpanosoma

1 (bases 1 to 599)

REFERENCE El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C., Gerrard,C., Leech,V., de Jong,P., Ullu,E., Melville,S., Donelson,J., Fraser,C. and Adams,M.

TITLE Use of BAC end sequences from Trypanosoma brucei GUTat 10.1 RPC1-93 Library for gene discovery and sequence-ready map construction

JOURNAL Unpublished (1999)

COMMENT Contact: Najib M, El-Sayed

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: nelsayed@tigr.org

Copies and high density filters may be purchased from BACPAC

Resources (<http://bacpac.med.buffalo.edu>). BAC end sequences search

page: <http://www.tigr.org/tcb/mdb/tbdb/>.

Seq primer: T7

Class: BAC ends.

FEATURES Location/Qualifiers

Source 1..599

/organism="Trypanosoma brucei"

/mol_type="genomic DNA"

/strain="TREU927/GUTat 10.1"

/db_xref="taxon:5691"

/clone="RPC193-DpnII-28M12"

/clone lib="RPC193-DpnII"

/note="Vector: PBAC3.6; Site 1: Bam HI; Site 2: Bam HI;

Constructed for The Institute for Genomic Research by Bohui Zhao in Pieter de Jong's laboratory (Roswell Park Cancer Institute, Buffalo, NY). Briefly, Trypanosoma brucei TREU927/GUTat 10.1 agarose embedded DNA was partially digested with a combination of Eco RI and Eco RI methylase (RPC193-EcoRI segment) or Dpn II (RPC193-DpnII segment). High molecular weight fragments were ligated in PBAC3.6 vector digested with Eco RI or Bam HI. Total coverage (both segments) > 90 X the haploid non-minichromosomal genome."

ORIGIN

Query Match 66.7%; Score 18; DB 8; Length 599;

Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: <http://genome.arizona.edu>
 PCR Primers
 FORWARD: gta aaa cga cgg cca gtg
 BACKWARD: gga aac aat gac cat g
 Plate: 16 row: G column: 03
 Seq primer: gta aaa cga cgg cca gtg.
 FEATURES
 source
 1. .719
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="Taxon:39947"
 /clone="OSNNEbi6G03"
 /tissue_type="leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lab="OSNNEB"
 /note="Vector: pBluescript II KS+; Site_1: EcoRI; Site_2:
 XbaI; 24 hrs after inoculation with Rice Blast (Che
 86061)"

ORIGIN

Query Match Score 18; DB 6; Length 719;
 Best Local Similarity 80.8%; Pred. No. 9.1e+02;
 Matches 21; Conservative 0; Mismatches 5;
 Indels 0; Gaps 0;
 Qy 1 GCGGCCGAATTGAAAGTGTATCC 26
 Db 246 GCAGCCGATTCAGTGTATCC 271

Search completed: January 7, 2005, 12:37:39
 Job time : 1617.95 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: January 7, 2005, 10:58:39 ; Search time 228.421 Seconds

(without alignments)

Title: US-09-786-502A-8

Pefect score: 30

Sequence: 1 tcggatctgtcaaggaggataggctgc 30

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters:

8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :	Score	Query	Match	Length	DB	ID	Description
	1	N_Geneseq_23Sep04:*					Aaa10271 Human CD2
	2	geneseqm1980:*					Aba92025 Human CD2
	3	geneseqm1990:*					Adk98533 Human imm
	4	geneseqn2001ab:*					Aat27650 CD2-CD28
	5	geneseqn2001bs:*					Aax90349 Human CD2
	6	geneseqn2002ab:*					Aaq21167 Human CD2
	7	geneseqn2002bs:*					Aat36257 Human CD2
	8	geneseqn2003ab:*					Aat14707 Human CD2
	9	geneseqn2003bs:*					Aav81203 Human CD2
	10	geneseqn2003cs:*					Aaa50581 Human cel
	11	geneseqn2003ds:*					Aaz229325 Human CD2
	12	geneseqn2004ab:*					Aab03175 Human lym
	1	30	100.0	30	3	AAA10271	Aaa10271 Human CD2
c	2	30	100.0	30	6	ABA22025	Aba92025 Human CD2
c	3	20.4	68.0	454	12	ADK98533	Adk98533 Human imm
c	4	20.4	68.0	875	2	AAT27650	Aat27650 CD2-CD28
c	5	20.4	68.0	1064	2	AAX90349	Aax90349 Human CD2
c	6	20.4	68.0	1514	2	AAQ21167	Aaq21167 Human CD2
c	7	20.4	68.0	1514	2	AAT16257	Aat36257 Human CD2
c	8	20.4	68.0	1514	2	AAT14707	Aat14707 Human CD2
c	9	20.4	68.0	1514	2	AAV81203	Aav81203 Human CD2
c	10	20.4	68.0	1514	2	AAA50581	Aaa50581 Human cel
c	11	20.4	68.0	1514	3	AAA50581	Aaa50581 Human cel
c	12	20.4	68.0	1514	3	AAZ229325	Aaz229325 Human CD2
c	13	20.4	68.0	1514	4	AAB03175	Aab03175 Human lym
c	14	20.4	68.0	1514	6	ABA9035	Aba9035 Human CD2
c	15	20.4	68.0	1514	10	ADD25537	Add25537 Binding d
c	16	20.4	68.0	1514	11	AD1319937	Adi1319937 Human CD2
c	17	20.4	68.0	1514	12	ADQ49343	Adq49343 Human CD2
c	18	20.4	68.0	1574	1	AAN90607	Aan90607 CD28 ant1
c	19	20.4	68.0	2405	2	AAT27652	Aat27652 LFA-1-be
c	20	20.4	68.0	3585	2	AAT27651	Aat27651 LFA-1-alp
c	21	20.4	68.0	3803	6	AAL49546	Aal49546 Human CD2

ALIGNMENTS

RESULT 1	ID	NAME	TYPE	VERSION	SEQUENCE	DESCRIPTION
AAA10271	AAA10271	standard; DNA; 30 BP.	XX	XX	AC	AAA10271;
			XX	XX	DT	03-JUL-2000 (first entry)
			XX	XX	DE	Human CD28 fragment downstream PCR primer.
			XX	XX	DE	CD28 cytoplasmic domain; prostate-specific membrane antigen; PSMA; J591 hybridoma; monoclonal antibody; single chain variable region; scFv; single chain antibody; fusion receptor; immune response; prostate cancer; PCR primer; ss.
			XX	XX	OS	Homo sapiens.
			XX	XX	PN	WO200014257-A1.
			XX	XX	PD	16-MAR-2000.
			XX	XX	PF	03-SEP-1999; 99WO-US020349.
			XX	XX	PR	04-SEP-1998; 98US-0099138P.
			XX	XX	PA	(SLOK) SLOAN KETTERING INST CANCER RES.
			XX	XX	PI	Sadelain M, Bander NH, Gong M, WPI; 2000-25702/22.
			XX	XX	PT	A fusion receptor composition having the structure:prostate-specific membrane antigen-single chain variable fragment:cytoplasmic domain, useful for treatment of cancer.
			XX	XX	PT	The invention relates to a novel fusion receptor composition having the structure:PSMA (prostate-specific membrane antigen)-scFv (single chain variable fragment):optional connector:cytoplasmic domain, where the fusion receptor is effective when expressed in a T-cell to promote a cellular immune response to PSMA. The PSMA-scFv is a single-chain antibody cloned from the V region genes of a hybrdoma specific for PSMA, such as J591. The optional connector is provided to give a spacing

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	30	100.0	30	3	AAA10271	Aaa10271 Human CD2	
c	2	30	100.0	30	6	ABA22025	Aba92025 Human CD2
c	3	20.4	68.0	454	12	ADK98533	Adk98533 Human imm
c	4	20.4	68.0	875	2	AAT27650	Aat27650 CD2-CD28
c	5	20.4	68.0	1064	2	AAX90349	Aax90349 Human CD2
c	6	20.4	68.0	1514	2	AAQ21167	Aaq21167 Human CD2
c	7	20.4	68.0	1514	2	AAT16257	Aat36257 Human CD2
c	8	20.4	68.0	1514	2	AAT14707	Aat14707 Human CD2
c	9	20.4	68.0	1514	2	AAV81203	Aav81203 Human CD2
c	10	20.4	68.0	1514	2	AAA50581	Aaa50581 Human cel
c	11	20.4	68.0	1514	3	AAA50581	Aaa50581 Human cel
c	12	20.4	68.0	1514	3	AAZ229325	Aaz229325 Human CD2
c	13	20.4	68.0	1514	4	AAB03175	Aab03175 Human lym
c	14	20.4	68.0	1514	6	ABA9035	Aba9035 Human CD2
c	15	20.4	68.0	1514	10	ADD25537	Add25537 Binding d
c	16	20.4	68.0	1514	11	AD1319937	Adi1319937 Human CD2
c	17	20.4	68.0	1514	12	ADQ49343	Adq49343 Human CD2
c	18	20.4	68.0	1574	1	AAN90607	Aan90607 CD28 ant1
c	19	20.4	68.0	2405	2	AAT27652	Aat27652 LFA-1-be
c	20	20.4	68.0	3585	2	AAT27651	Aat27651 LFA-1-alp
c	21	20.4	68.0	3803	6	AAL49546	Aal49546 Human CD2

CC between the PSMA-scFv and the cytoplasmic domain, such that both retain
 CC substantial function. The cytoplasmic domain directs the function of the
 CC fusion receptor and is generally the cytoplasmic domain of a molecule
 CC which functions as a transducer of a mammalian immune response by a molecule
 CC presence of an MHC (major histocompatibility complex)-peptide complex or
 CC costimulatory factor. Examples of cytoplasmic domains that may be
 CC employed in the present invention include the T-cell receptor gamma-chain
 CC cytoplasmic domain and the CD28 cytoplasmic domain. In a method of the
 CC invention, an expression vector encoding the fusion receptor is
 CC transduced into primary T-lymphocytes obtained from the patient to be
 CC treated. The transduced lymphocytes are returned to the patient where
 CC they secrete interleukin-2 (IL-2) and proliferate in response to PSMA-
 CC positive cells. The resulting cytotoxic lymphocytes specifically lyse
 CC cells expressing PSMA and can thus be used to target PSMA-positive tumour
 CC cells. The fusion receptor promotes a cellular immune response to PSMA
 CC and is useful for the treatment of prostate cancer and other cancers that
 CC express PSMA. Sequences AAA0270-A10271 represent PCR primers used in an
 CC exemplification of the present invention to amplify cDNA encoding the
 CC cytoplasmic and transmembrane domains and part of the extracellular
 CC domain of human CR28. This was used to construct a gene encoding a fusion
 CC receptor comprising the PSMA-scFv and the CD28 fragment

XX Sequence 30 BP; 6 A; 6 C; 11 G; 7 T; 0 U; 0 Other;

Query Match 100.0% Score 30; DB 3; Length 30;
 Best Local Similarity 100.0% Pred. No. 0.00094;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCGAGGATCTTGTCAGGACGATAGGCC 30

Db 1 TCGAGGATCTTGTCAAGACGATAGGCC 30

RESULT 2
 AB92025
 ID AB92025 standard, DNA, 30 BP.
 XX AC
 XX 23-MAY-2002 (First entry)

XX DE Human CR28 cDNA downstream PCR primer.

KW T cell; receptor; human; antitumour; immunostimulant; cancer;
 KW therapy; neuroblastoma; melanoma; sarcoma; small lung carcinoma;
 KW brain tumour; disialoganglioside GD2; imaging; PCR; primer; ss.
 XX Homo sapiens.
 OS US2002018783-A1.
 PN 14 FEB-2002.
 XX PR 30-SEP-1997; 97US-00940544.
 XX PR 20-MAR-1997; 97WO-US004422.XX PA (SDDE/) Sadelain M.
 PA (CHEU/) CHEUNG N V.
 PA (KRAU/) KRAUSE A.
 PA (GUOH/) GUO H.
 XX PI Sadelain M, Cheung NV, Krause A, Guo H;

XX DR; 2002-239251/29.

XX PT New fusion protein, useful for inducing host immune response, comprises

PT variable region of light chain of an antibody linked to variable region

PT of antibody, CD28 receptor signalling domain and transmembrane domain.

XX Example 3; Page 4; 9pp; English.

XX The present sequence is that of a downstream primer, used with the

CC upstream primer given in AB92024, in the PCR amplification of a segment
 CC of human T cell surface receptor CD28 cDNA. The amplified segment encodes
 CC part of the extracellular domain of CD28, and the transmembrane and the
 CC cytoplasmic domains. Plasmid pbsCD28 was used as template. The 5' primer
 CC contains an Not I site, while the 3' primer contains a BamHI site, and
 CC these sites in the PCR product were used to facilitate insertion into
 CC retroviral vector SFG. A cell-surface molecule capable of CD28 signalling
 CC in T cells interacting with disialoganglioside GD2 tumours was
 CC constructed. This comprised the antigen-binding site of a GD2-specific
 CC antibody and the transmembrane and signalling domains of the invention
 CC molecule. This is an example of fusion proteins of the invention
 CC comprising a single chain Fv antibody linked to CD28 receptor. T cells
 CC expressing such fusion proteins exhibit enhanced survival when
 CC reintroduced to an in vivo environment. They can be used to induce an
 CC immune response to cells, particularly tumour cells which express the
 CC antigen to which the anti-GD2 scFv are specific. Cells expressing a fusion
 CC protein including an anti-GD2 scFv are useful for treatment of melanomas,
 CC neuroblastomas, small lung carcinoma, sarcomas and brain tumours that
 CC express GD2 as a surface antigen. Cells expressing the fusion proteins of
 CC the invention can also be used for in vitro purging of stem cells or bone
 CC marrow and for in vivo targeting of tumour cells and other antigen-
 CC bearing cells for imaging

XX SQ Sequence 30 BP; 6 A; 6 C; 11 G; 7 T; 0 U; 0 Other;
 XX Query Match 100.0% Score 30; DB 6; Length 30;
 Best Local Similarity 100.0% Pred. No. 0.00094;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1 TCGAGGATCTTGTCAGGACGATAGGCC 30
Db	1 TCGAGGATCTTGTCAAGACGATAGGCC 30

RESULT 3

ADK98583/C
 ID ADK98583 standard, cdNA; 454 BP.
 XX DT 03-JUN-2004 (first entry)

XX XX Human immune response associated protein IRAP-11 cDNA.
 XX XX immune response associated protein; IRAP; antiarteriosclerotic;
 KW KW cytostatic; neuroprotective; anticardiovascular; hepatotropic;
 KW KW cerebroprotective; antiinflammatory; nootropic; vasotrophic;
 KW KW arteriosclerosis; cirrhosis; cancer; stroke; Alzheimer's disease;
 KW Parkinson's; Crohn's; gene therapy; human; ss; gene.
 XX Homo sapiens.
 OS XX OS WO2004020593-A2.

XX XX PD 11-MAR-2004.

XX XX PF 26-AUG-2003; 2003WO-US026988.
 XX XX PR 30-AUG-2002; 2002US-0407561P.
 XX XX PR 11-SEP-2002; 2002US-0410178P.
 XX XX PR 13-SEP-2002; 2002US-0410571P.
 XX XX PR 18-OCT-2002; 2002US-0419061P.
 XX XX PR 25-OCT-2002; 2002US-0421445P.
 XX XX PA (INCYT) INCYTE CORP.

XX XX PI Ramkumar J, Swarnakar A, Elliott VS, Hafalia AJA, Richardson TW,
 PI Lee SY, Lindquist EA, Marquis JP, Chawla NK, Khare R, Becha SD;
 XX DR WPI; 2004-239178/22.
 DR N-PADB; ADK98548.
 XX PT New isolated immune response associated proteins (IRAP) polypeptide and

PT polynucleotide, useful for diagnosing and/or treating disorders with aberrant expression of IRAP, such as arteriosclerosis, cirrhosis, cancer and stroke.

PT Claim 5; SEQ ID NO 46; 207pp; English.

XX The invention relates to a novel isolated immune response associated protein (IRAP) comprising any of 35 fully defined sequences given in the specification. The polypeptide of the invention demonstrates antiangiogenic, cytoprotective, neuroprotective, anticancer, stroke, hepatotropic, cerebroprotective, antiinflammatory, nociceptive and vasotropics activities and may be useful for treating a disease or condition associated with decreased expression or overexpression of functional immune response associated proteins, while the antibody is useful for diagnosing a condition or disease associated with the expression of IRAP, such as arteriosclerosis, cirrhosis, cancer, stroke, Alzheimer's disease, Parkinson's disease and Crohn's disease. Furthermore, the molecules of the invention may be utilised during gene therapy procedures. The current sequence is that of a human IRAP cDNA of the invention.

XX Sequence 454 BP; 109 A; 142 C; 105 G; 98 T; 0 U; 0 Other;

CC Best Local Similarity 95.5%; Pred. No. 31;

CC Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CC Qy 9 CGTGTCAAGGAGCATAGCTGC 30

CC Db 391 CGTGTCAAGGAGCATAGCTGC 370

RESULT 4
AAT27650/c
ID AAT27650 Standard; cDNA; 875 BP.

XX AAT27650;
AC
DT 24-AUG-1996 (first entry)

XX CD2-CD28 chimeric receptor.

XX Human; chimeric receptor; CD2; extracellular domain; CD28;

XX transmembrane region; intracellular domain; retrovirus; vector;

XX packaging cell culture; gene transfer; LFA-3; Selective activation;

XX cytotoxic T-lymphocyte; antitumour; immunotherapy; cancer; therapy;

XX virucide; ss.

XX Homo sapiens.

XX WO9613584-A1.

XX PN 09-MAY-1996.

XX PP 01-NOV-1995; 95WO-US014171.

XX PR 01-NOV-1994; 94US-00332993.

XX 07-JUN-1995; 95US-00480577.

XX (TARG-) TARGETED GENETICS CORP.

XX PI Feldhaus AL, Jones LA;

XX PD 09-MAY-1997/24.

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY24470.

XX XX PT Oligo:nucleotide which reduces CD28 gene expression in T cells - for

XX PT treating immune system diseases, e.g. graft vs. host disease, septic shock, psoriasis, etc.

XX PS Disclosure; Fig 1C; 45pp; English.

XX XX Tam RC;

XX XX WPI; 1996-384228/38.

XX DR P-PSDB; AAY244

Query Match 68.0%; Score 20.4; DB 2; Length 1064;
 Best Local Similarity 95.5%; Pred. No. 36; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 1;

RESULT 7
 AAT36257/C
 ID AAT36257 standard; cDNA; 1514 BP.
 XX
 AC AAT36257;
 AC
 XX
 DT 25-MAR-2003 (revised)
 DT 15-APR-1997 (first entry)

RESULT 6
 AAC21167/C
 ID AAC21167 standard; DNA; 1514 BP.
 XX
 AC AAC21167;
 XX
 DT 25-MAR-2003 (revised)
 DT 21-MAY-1992 (first entry)

Human CD28 antigen coding sequence.

KW cloning technique; cell surface antigen; immunodiagnosis; tumour; ss.

XX Homo sapiens.

OS Homo sapiens.

XX

Key Location/Qualifiers
 FH 100..762
 FT /*tag= a
 FT /product= "CD28"
 FT 100..153
 FT /*tag= b
 PN WO9201049-A.
 XX
 PD 23-JAN-1992.
 PP 13-JUL-1990; 90US-00553759.
 XX
 PR 13-JUL-1990; 90US-00553759.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Seed B, Aruffo A, Amiot M;
 XX
 DR WPI:1992-056864/07.
 DR P-PSDB; AAR20805.
 XX
 PT New CD53 cell surface antigen and DNA encoding it - for immuno-therapy
 PT and diagnosis of haematopoietic neoplasms, etc.

Example 3: Fig 7; 160pp; English.

XX A library of recombinants having inserts greater than 0.8kb in size was prepared in the pBHM vector (see AA021166), from 1 microgram of polyA (plus) RNA isolated from the human lymphoblastoid cell line JY. The library was screened for CD28 cDNA clones which were isolated using the antibody enrichment method (see e.g. AA021164 or AA021165). After the third transfaction, COS cells were panned with a specific anti-CD28 Ab. A Hirt supernatant was generated and transformed into E.coli. DNA was prepared from the resulting colonies and transfected into COS cells. Surface expression of CD28 antigen was detected in 3 of the 8 transfected cultures by indirect immunofluorescence. The cDNA insert from one of the positive clones was sequenced. (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;

Query Match 68.0%; Score 20.4; DB 2; Length 1514;
 Best Local Similarity 95.5%; Pred. No. 37; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 1;

Qy 9 CTTGTCAAGGAGCATAGGGTGC 30
 DB 766 CGTGTCAAGGAGCATAGGGTGC 745

RESULT 8
 AAT14707/C
 ID AAT14707 standard; cDNA; 1514 BP.

DT 10-MAY-1999 (first entry)
 XX Human CD28 cDNA.
 XX CD28; cell surface antigen; human; T cell antigen; T lymphocyte;
 KW cDNA library; ss.
 XX OS *Homo sapiens*.
 XX Key Location/Qualifiers
 CDS 100..762 /*tag= a
 sig_peptide 100..153 /*tag= b
 mat_peptide 154..759 /*tag= c
 XX FT Homo sapiens.
 XX Key Location/Qualifiers
 CDS 100..762 /*tag= a
 sig_peptide 100..153 /*tag= b
 mat_peptide 154..759 /*tag= c
 XX FT CDS
 PR 07-JUN-1995; 95US-00485447.
 PR 25-FEB-1988; 88US-00160416.
 PR 13-JUL-1989; 89US-00379076.
 PR 23-MAR-1990; 90US-00498809.
 PR 13-JUL-1990; 90US-00553759.
 PR 01-DEC-1992; 92US-00983647.
 XX PR 25-FEB-1988;
 PR 13-JUL-1989;
 PR 23-MAR-1990;
 PR 13-JUL-1990;
 PR 01-DEC-1992;
 XX PR 88US-00160416.
 PR 90US-00379076.
 PR 90US-00498809.
 PR 90US-00553759.
 PR 92US-00983647.
 XX PR (GEHO) GEN HOSPITAL CORP.
 PA P-PSDB; AAW88451.
 XX PT cDNA encoding human Cd40 antigen - useful for cloning cDNA encoding cell surface antigens, constructing cDNA libraries, expressing vectors for expression in eukaryotic cells or their fragments.
 XX PS (GEHO) GEN HOSPITAL CORP.
 XX CC This nucleotide sequence comprises human CD28 cDNA. The cDNA was isolated from HPB-ALL human T-cell tumour cells using a novel method for cloning cDNAs from mammalian expression libraries. The method is based on transient expression of an antigen in eukaryotic cells and physical selection of cells expressing the antigen by adhesion to an antibody-coated substrate. The method is useful for the isolation and molecular cloning of any protein which can be expressed and transported to the cell surface membrane of a eukaryotic cell. It has been used to clone genes (see AAW81198-2201) encoding cell surface antigens such as CD1a, CD1b, CD1c, CD2, CD6, CD7, CD13, CD14, CD16, CD19, CD20, CD22, CD26, CD27, CD28, CD31, CD32, CD32b, CD33, CD34, CD36, CD37, CD38, CD39, CD40, CD43, CD44, CD53, ICAM, LFA-3, FcR γ , TlIISa and Leu8 (see AAW86188-62). AAW89151-52 and AAW88451 (see AAV81198) is specifically claimed. CD28 polypeptide (see AAW88451) has been expressed in transfected COS cells
 XX SQ Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;
 SQ Query Match 9 CTGTCTCAGGAGCGATAGGGTGC 30
 Best Local Similarity 68.0%; Score 20.4%; DB 2; Length 1514;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 9 CTGTCTCAGGAGCGATAGGGTGC 30
 Db 766 CGTGTCTCAGGAGCGATAGGGTGC 745
 RESULT 11
 AA50581/C
 ID AA50581 standard cDNA; 1514 BP.

XX AC AAA50581;
 XX DT 19-DEC-2000 (first entry)
 XX DB Human cell surface antigen CD28 cDNA.
 XX KW CD28; cell surface antigen; human; immunoselection; panning; gene therapy; gene diagnosis; immunotherapy; gene therapy; immune disorder; infection; asthma; immune-complex disease; amyloidosis; multiple sclerosis; parasitic disease; ss.
 XX OS Homo sapiens.
 XX Key Location/Qualifiers
 CDS 100..762 /*tag= a
 PR 29-AUG-2000.
 XX PR 28-OCT-1998; 98US-00181612.
 XX PR 25-FEB-1988;
 PR 13-JUL-1989;
 PR 23-MAR-1990;
 PR 13-JUL-1990;
 PR 01-DEC-1992;
 XX PR (GEHO) GEN HOSPITAL CORP.
 PA P-PSDB; AAY96128.
 XX PT Isolated nucleic acid molecule encoding the CD19 cell surface antigen, useful for immunodiagnosis and immunotherapy of immune-mediated infections or disorders, e.g. asthma, immune-complex disease, parasitic diseases.
 XX PS Example 3; Fig 7A-B; 75pp; English.
 XX CC The present sequence is that of cDNA encoding human cell surface antigen (CSA) CB28 (see AAY96128). The cDNA was derived from a human T-cell tumour line HPB-ALL cDNA library using a new method for cloning CSA cDNAs. The method is based upon transient expression of CSA in eukaryotic cells and physical selection of cells expressing the antigen by adhesion to (panning on) an antibody-coated substrate such as a culture dish. The predicted amino acid sequence of CB28 suggests an integral membrane protein with a single membrane-spanning hydrophobic domain terminating in a 41-amino acid cytoplasmic domain. CSA nucleic acids isolated by the method of the invention, and the proteins they encode, are useful for immunodiagnostic and immunotherapeutic applications, including the diagnosis and treatment of immune-mediated infections, diseases, and disorders in animals, including humans. These disorders include asthma, immune-complex disease, amyloidosis, parasitic diseases or multiple sclerosis.
 XX SQ Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;
 SQ Query Match 9 CTGTCTCAGGAGCGATAGGGTGC 30
 Best Local Similarity 68.0%; Score 20.4%; DB 3; Length 1514;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 9 CTGTCTCAGGAGCGATAGGGTGC 30
 Db 766 CGTGTCTCAGGAGCGATAGGGTGC 745

RESULT 12	ID AA229325/C	AA229325 standard; cDNA; 1514 BP.	ID AAS03175 standard; cDNA; 1514 BP.	AAS03175;
XX	XX	XX	XX	29-AUG-2001 (first entry)
AC	AC	Human lymphocyte cell surface antigen CD28 cDNA sequence.	DE	
XX	XX	Human; lymphocyte cell surface antigen; immune-mediated disease; CD28; infection; immune deficiency disorder; hypersensitivity; inflammation; systemic lupus erythematosus; platelet disorder; rheumatoid arthritis; transplant rejection; asthma; ss.	KW	
DT 29-FEB-2000 (first entry)	DE Human CD28 receptor cDNA.	XX	KW	
XX	XX	Major histocompatibility complex; MHC classII; proton motor force; mitochondrial membrane potential; mitochondrial metabolism; cancer; autoimmune disease; neurodegenerative disorder; ds.	KW	
XX	XX	Homo sapiens.	OS Homo sapiens.	
OS	OS	Location/Qualifiers	Key	XX
XX	XX	100..762	CDS	100..762
FH	FH	/*tag= a	FT	/*tag= a
PT CDS	PT	/product= "CD28 antigen"	FT	/product= "CD28 antigen"
PT	PT	100..153	FT	100..153
FT	FT	/*tag= b	FT	/*tag= b
XX	XX	154..759	FT	154..759
PN W099393-A2.	PN	/*tag= c	XX	/*tag= c
XX	XX	US6218525-B1.	PN	
XX	XX	17-APR-2001.	PD	
PD 28-OCT-1999.	XX	92US-00983647.	XX	
XX	XX	01-DEC-1992;	PF	
PR 30-MAR-1999;	PR	99WO-US006874.	XX	
XX	XX	88US-00160416.	PR	
PR 17-APR-1998;	PR	98US-00822509.	PR	
PR 29-JUL-1998;	PR	98US-0094515P.	PR	
PR 24-SEP-1998;	PR	98US-0101580D.	PR	
XX	XX	90US-00553759.	XX	
(UYVE-) UNIV VERMONT.	PA (GEHO) GEN HOSPITAL CORP.		PA	
XX	XX		Seed B, Aruffo A, Simmons D;	
XX	XX		WPI; 2001-289848/30.	
DR DR	DR		DR P-PSDB; AAU02437.	
DR P-PSDB; AAU4294.	XX			
XX	XX			
PT Use of cell surface and membrane characteristics for developing products for treating cancers, autoimmune diseases or neurodegenerative diseases.	PT New recombinant DNA encoding CD28 useful for diagnosing and treating immune-mediated diseases, infections or disorders, e.g. systemic lupus erythematosus, asthma, transplant rejection, rheumatoid arthritis.			
XX	XX			
PS Disclosure; Page 120-121; 123pp; English.	PS Example 3; Fig 7A-7B; 72pp; English.			
XX	XX			
CC The present sequence encodes human CD28 receptor. CD28 is a homodimeric glycoprotein involved in a secondary signalling pathway in the activation of T-cell proliferation. B7 on nerve cells can interact with CD28 on the immune cell leading to immune cell activation. The regulation of cell surface expression of MHC classII and co-stimulatory molecule B7 can be manipulated by regulating the intracellular dissipation of proton motor force which can be assessed in terms of mitochondrial membrane potential. These methods can be used for regulating cell growth and division to control disease processes by manipulating mitochondrial metabolism and the expression of cell surface immune proteins. They can be used for treating diseases associated with excessive cellular division, aberrant differentiation, and premature cellular death, e.g. cancers, autoimmune diseases, neurodegenerative disorders etc.	CC The present sequence encoding for human lymphocyte cell surface antigen CD28 is used to obtain a new genetically engineered cDNA sequence encoding the CD28 amino acid extracellular domain sequence (amino acids 1 CC -134 given in AU02437) and/or comprising nucleotides 100-759, 154-555 or 154-759 of the CD28 cDNA sequence (AU02437, AAS03175-AAS03195) are CC antigen cDNA sequences (AU03172, AAS03173-AAS03195) are described in the present invention. The invention relates to a novel CC method of cloning cDNA encoding cell surface antigens and efficient CC construction of cDNA libraries. Also described are 2 expression vectors CC (AAS03171, AAS03174) which provide high level expression in eukaryotic CC host cells. The purified genes and proteins are useful for CC immunodiagnostic and immunotherapeutic applications, such as in the CC diagnosis and treatment of immune-mediated diseases, infections or CC diseases, diseases of immediate type of hypersensitivity, asthma, CC hypersensitivity pneumonitis, systemic lupus erythematosus, rheumatoid CC arthritis, acute and chronic inflammation, platelet disorders, plasma and CC other cell neoplasms, parasitic diseases, multiple sclerosis, Guillain- Barre syndrome and tissue and organ transplant rejection. The sequences CC can also be used to identify, isolate and purify other antibodies and CC antigens			
XX	XX			
SQ Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;	Query Match 68.0%; Score 20.4; DB 3; Length 1514;			
XX	XX			
Best Local Similarity 95.5%; Pred. No. 37; Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	Best Local Similarity 95.5%; Pred. No. 37; Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy 9 CTGTCAGGAGGTAGGGTC 30	Qy 9 CTGTCAGGAGGTAGGGTC 30			
Db 766 CGGTCAGGAGGTAGGGTC 745	Db 766 CGGTCAGGAGGTAGGGTC 745			
SQ Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;	Query Match 68.0%; Score 20.4; DB 4; Length 1514;			
Best Local Similarity 95.5%; Pred. No. 37;	Best Local Similarity 95.5%; Pred. No. 37;			

Matches	21;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;	Db	766 CGTGTAGGAGCGATGGCTGC 745
QY	9	CTTGTCAGGAGCGATAGGGCTGC	30							RESULT 15	
Db	766	CGTGTAGGAGCGATAGGGCTGC	745							ID ADD25537/C	
										ID ADD25537 standard; DNA; 1514 BP.	
RESULT 14										XX	
AB99031/C										AC ADD25537;	
ID AB99035 standard; DNA; 1514 BP.										XX	
XX										XX DT 15-JUN-2004 (first entry)	
AC										DE Binding domain-immunoglobulin fusion protein-associated DNA #53.	
ABA99035;										XX	
XX										ds; Binding domain; immunoglobulin; fusion protein; cyostatic; antiarthritic; immunosuppressive; antidiabetic; antithyroid;	
XX										KW neuroprotective; hinge region; immunoglobulin heavy chain;	
DE Human CD28 gene.										KW CH2 constant region; CH3 constant region; IgG1; antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation; malignant condition; B-cell disorder; melanoma; carcinoma; rheumatoid arthritis; myasthenia gravis; Grave's disease; type I diabetes mellitus; multiple sclerosis; autoimmune disease; polyarteritis nodosa; vasculitis; systemic lupus erythematosus; Sjogren's syndrome; psoriasis; dermatomyositis; polymyositis; sepsis; gene; ds.	
XX										KW Unidentified.	
Human CD28										OS US2003118592-A1.	
XX										PN XX	
marrow transplant rejection; organ transplant rejection; psoriasis; tissue rejection; autoimmune disease; diabetes mellitus; multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus; systemic sclerosis; dermatomyositis; polymyositis; Sjogren's syndrome; polyarteritis nodosa; vasculitis; gene; ds.										PD 26-JUN-2003.	
XX										XX 25-JUL-2002; 2002US-00207655.	
Human; CD28; immune response; T cell; graft-versus-host disease; GVHD; marrow transplant rejection; organ transplant rejection; psoriasis; tissue rejection; autoimmune disease; diabetes mellitus; multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus; systemic sclerosis; dermatomyositis; polymyositis; Sjogren's syndrome; polyarteritis nodosa; vasculitis; gene; ds.										XX 17-JAN-2001; 2001US-036758P.	
XX										PR 17-JAN-2002; 2002US-0005530.	
DE Human CD28 gene.										PR 03-JUN-2002; 2002US-0385691P.	
XX										XX (GENE-) GENECRAFT INC.	
XX										XX PA	
XX										PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;	
XX										XX DR WPI; 2003-801317/75.	
XX										XX New binding domain-immunoglobulin fusion protein, useful for treating a subject having or suspected of having a malignant condition or a B-cell disorder, e.g. melanoma, Grave's disease or autoimmune disease.	
XX										XX Disclosure; SEQ ID NO 98; 157PP; English.	
XX										XX Unidentified	
XX										CC XX Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;	
XX										CC Query Match Score 20.4%; Best Local Similarity 95.5%; Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
XX										CC DB 9 CTTGTCAGGAGCGATGGCTGC 30	
XX										RESULT 16	
XX										AD131990/C	
XX										AD131990 standard; cDNA; 1514 BP.	
XX										DB Human cDNA #1316.	
XX										DR Human; gene; ss; immunological response; immunopathological condition; Crohn's disease; asthma; ulcerative colitis; hypereosinophilia; irritable bowel syndrome; osteoarthritis; rheumatoid arthritis; acute monocytic leukaemia; antiasthmatic; antiulcer;	
XX										SQ Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;	
XX										Query Match Score 20.4%; Best Local Similarity 95.5%; Pred. No. 37; Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
XX										DB 766 CGTGTAGGAGCGATAGGGCTGC 30	
XX										QY 9 CTTGTCAGGAGCGATAGGGCTGC 30	

KW osteopathic; antiarthritic; antirheumatic; cytostatic.
 XX KW cell surface antigen; immune-mediated disorder; asthma;
 OS KW rheumatoid arthritis; multiple sclerosis; vasculitis; inflammation; 88;
 Homo sapiens.
 XX KW gene; human.
 PN XX
 US6607879-B1.
 XX OS
 PD XX Homo sapiens.
 19-AUG-2003.
 XX OS
 XX US2004072283-A1.
 XX OS
 PF XX
 09-FEB-1998; 98US-00023655.
 XX PR 15-APR-2004.
 XX PR 17-APR-2001; 2001US-00836544.
 XX PR 25-FEB-1988; 88US-00160416.
 PA PR 13-JUL-1989; 89US-00379076.
 (INCY-) PA PR 23-MAR-1990; 90US-00498009.
 PA PR 13-JUL-1990; 90US-00553759.
 PA PR 01-DEC-1992; 92US-00983647.
 XX PA (SEED/) SEED B.
 PA (ALLEL) ALLEN J.
 PA (ARUFFO) ARUFFO A.
 PA (CAMEINI) CAMEINI D.
 PA (LAUFER) LAUFER L.
 PA (OQUENDO) OQUENDO C.
 PA (SIMMONS) SIMMONS D.
 PA (STAMENKOVIĆ) STAMENKOVIĆ I.
 PA (STENGELIN) STENGELIN S.
 PA (AMIOU) AMIOU M.
 XX PI Seed B, Allen J, Aruffo A, Camerini D, Lauffer L, Oquendo C;
 PI Simmons D, Stamenkovic I, Stengelin S, Amiot M;
 XX DR WPI; 2004-228571/30.
 XX DR P-PSDB; ADO49344.
 XX New cloning cDNA segments encoding cell surface antigens of human
 PT lymphocytes, useful in diagnosing and treating asthma, rheumatoid
 PT arthritis, multiple sclerosis, vasculitis and inflammation and
 PT infections.
 XX Example 3; Fig 7; 75PP; English.
 PS CC The invention relates to a composition comprising a plurality of cDNAs
 for detecting the altered expression of genes in an immunological
 response. The invention also relates to a method of diagnosing or
 monitoring the treatment of an immunopathological condition in a sample,
 comprising obtaining nucleic acids from a sample, contacting the nucleic
 acids of the sample with an array comprising the plurality of cDNAs under
 conditions to form one or more hybridisation complexes, detecting the
 hybridisation complexes and comparing the levels of the detected
 hybridisation complexes with the level of hybridisation complexes
 detected in a non-diseased sample, where an altered level of the detected
 hybridisation complexes correlates with the presence of an
 immunopathological condition. Also disclosed are an expression profile
 comprising a microarray and a plurality of detectable complexes and a
 method for identifying a plurality of polynucleotide probes. The cDNAs
 are useful as hybridisable array elements in a microarray for monitoring
 the expression of target polynucleotides. The microarray can be used in
 the diagnosis of an immunopathology, such as Crohn's disease, asthma,
 ulcerative colitis, hyperimmunophilia, irritable bowel syndrome,
 osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in
 identifying agents for the treatment of the diseases. The microarray may
 also be used in drug discovery and development, toxicological and
 carcinogenicity studies, forensics or pharmacogenomics. The composition
 may also be used in purification of a subpopulation of mRNAs cDNAs or
 genomic fragments. This sequence represents a human cDNA of the
 invention. Note: The sequence data for this patent did not form part of
 the printed specification but was obtained in electronic format directly
 from USPTO at seqpto.uspto.gov/sequence.html.
 XX SQ Sequence 1514 BP; 404 A; 360 C; 337 G; 413 T; 0 U; 0 Other;
 XX SQ Query Match 68.0%; Score 20.4; DB 11; Length 1514;
 Best Local Similarity 95.5%; Pred. No. 37; Pred. No. 37;
 Matches 21; Conservative 0; Mismatches 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 9 CTGTCAGGAGCGATAGGGTGC 30
 Db 766 CGTGTCAAGAGCGATAGGTGC 745
 RESULT 17
 ADD49343/C
 ID ADD49343 Standard; cDNA; 1514 BP.
 XX AC AAN90607 standard; cDNA; 1574 BP.
 AC AC
 AC AAN90607;
 XX DT 09-SEP-2004 (revised)
 DT 25-MAR-2003 (revised)
 DT 31-OCT-2002 (revised)
 DE Human CD28 antigen cDNA.

RESULT 18
 AAN90607/C
 ID AAN90607 standard; cDNA; 1574 BP.
 XX AC AAN90607;
 AC AC
 AC AAN90607;
 XX DT 09-SEP-2004 (first entry)
 XX DT 15-JUL-2004 (first entry)
 DE Human CD28 antigen cDNA.

DT	20-DEC-1989	(first entry)
XX	PN	WO9613584-A1.
DE	XX	
CD28 antigen cDNA.	XX	09-MAY-1996.
XX	PD	
Tyrosine suppressor gene; ASV LTR; human cytomegalovirus AD169 enhancer;	XX	09-MAY-1996.
HIV box; immunoselection; immune deficiency diseases; vasculitis;	XX	01-NOV-1995;
systemic lupus erythematosus; rheumatoid arthritis; neoplasms; ss.	XX	PP
XX	PR	95WO-US014171.
OS	XX	
Homo sapiens.	PR	94US-00332993.
Unidentified.	PR	95US-00480577.
XX	XX	
KEY	PA	(TARG-) TARGETED GENETICS CORP.
CDS	XX	
FT	PI	Feldhaus AL, Jones LA;
FT	XX	
*tag= a	PS	Example 1; Page 57-58; 82pp; English.
sig_peptide	XX	
FT	DR	WPI; 1996-239497/24.
XX	XX	
Location/Qualifiers.	PT	TH independent cytotoxic T lymphocytes useful in tumour or viral therapy
100..819	-	- are activated by adhesion to antigen displaying cells.
EP330191-A.	XX	
XX	XX	
30-AUG-1989.	XX	
XX	CC	The sequence is an XbaI-SalI fragment encoding a chimeric receptor,
PP	CC	composed of the extracellular region (XR) of leukocyte function-
23-FEB-1989;	CC	associated antigen-1 beta-chain (LFA-1-beta), of the integrin family),
89EP-00103127.	CC	fused via the transmembrane region (TM) of CD28 to the intracellular
XX	CC	region (IC28) of CD28-TM and IC28 coding sequences (bp 556-767) are
(GEHO) GEN HOSPITAL CORP.	CC	amplified by PCR from human CD28 cDNA, and the XR coding sequence (bp 1-
PA	CC	2072) is amplified by PCR from human LFA-1-beta cDNA. The sequences are
XX	CC	cloned to form plasmid pSKLFA-1-beta/CD28, and subcloned in HyTK-CMV-Cat-
Seed B, Allen J, Aruffo A, Camerini D, Laufer L, Oquendo CP;	CC	to form HyTK-CMV-LFA-1-beta/CD28. This retrovirus vector is introduced
PI	CC	into a psi-2 ecotropic packaging cell culture to give a stable PA317 cell
Simmons D, Stamenkovi I;	CC	line producing high titers of infectious particles, which are then used
XX	CC	to infect T-lymphocytes. The resulting cytotoxic T-lymphocytes (CTLs) are
WPI; 1989-250302/35.	CC	selectively activatable, and have a lessened dependence on helper T-
P-PSDB; AAP91324.	CC	CC lymphocytes (TH cells), and/or growth factors. The chimeric receptor binds
Rapid immuno-selection cloning - used to clone genes encoding cell	CC	CC to the LFA-1 ligand (intracellular adhesion molecule-1, ICAM-1 or CD54)
surface antigens associated with mammalian T lymphocytes	CC	but IC28 transmits the proliferative signal of CD28. The CTLs mediate
Disclosure; Fig 7; 69pp; English.	CC	CC lysis of a target cell bearing a cognate antigen, and are useful in
XX	XX	CC directed cancer immunotherapy, and in treatment of viral infections
This is used for cloning into a vector which transforms COS cells. The	SQ	Sequence 2405 BP; 503 A; 736 C; 739 G; 427 T; 0 U; 0 Other;
vectors can be used to isolate any protein by immunoselection, and clones	Query	Query Match Score 20.4; DB 2; Length 2405;
are easy to manipulate. (Updated on 31-OCT-2002 to add missing OS field.)	Best Local Similarity 68.0%; Pred. No. 40;	
(Updated on 25-MAR-2003 to correct PA field.)	Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0	
CC	Qy	9 CTTGTCGGAGCGATAGGCTGC 30
CC	Db	2398 CCTGTCAAGGCGATAGGCTGC 2377
Revised record issued on 09-SEP-2004 : Correction to keywords	RESULT 20	
XX	ID	AAT27651/C
Sequence 1574 BP; 423 A; 375 C; 347 G; 429 T; 0 U; 0 Other;	XX	
XX	ID	AAT27651;
CC	AC	
Revised record issued on 09-SEP-2004 : Correction to keywords	XX	
XX	DT	24-AUG-1996 (first entry)
Sequence 1574 BP; 423 A; 375 C; 347 G; 429 T; 0 U; 0 Other;	XX	
XX	XX	LFA-1-beta-CD28 chimeric receptor.
CC	DE	LFA-1-alpha-CD28 chimeric receptor.
CC	XX	Human; chimeric receptor; LFA-1-beta; extracellular domain; CD28;
CC	XX	transmembrane region; intracellular domain; retrovirus; vector;
CC	XX	packaging cell culture; gene transfer; ICAM-1; selective activation;
CC	XX	cytotoxic T-lymphocyte; antitumour; immunotherapy; cancer; therapy;
CC	XX	viricide; ss.
CC	XX	Homo sapiens.
CC	OS	OS
CC	XX	WO9613584-A1.
CC	XX	09-MAY-1996.
OS	XX	

PF 01-NOV-1995; 95WO-US014171.
 XX
 PR 01-NOV-1994; 94US-00332993.
 PR 07-JUN-1995; 95US-00480577.
 XX (TARG-) TARGETED GENETICS CORP.
 PA Feldhaus AL; Jones LA;
 XX PI DR WPI; 1996-23949/7/24.
 XX PT TH independent cytotoxic T lymphocytes useful in tumour or viral therapy
 - are activated by adhesion to antigen displaying cells.
 XX PS Example 1: Page 55-57; 82PP; English.

XX The sequence is an XbaI-KpnI fragment encoding a chimeric receptor, composed of the extracellular region (XKR) of leukocyte function-associated antigen-1, alpha-chain (LFA-1-alpha), of the integrin family, fused via the transmembrane region (TM) of CD28 to the intracellular region (IC28) of CD28, TM and IC28 coding sequences (bp 556-767) are amplified by PCR from human CD28 cDNA, and the XKR coding sequence (bp 1-3358) is amplified by PCR from human LFA-1- alpha cDNA. The sequences are cloned to form plasmid pSKLFLA-1- alpha/CD28, and subcloned in tg1SpplusNeo-CMV to form tg1SpplusUBO- CMV-LFA-1-alpha/CD28. This retrovirus vector is introduced into a psi-2 ecotropic packaging cell culture to give a stable PA317 cell line producing high titers of infectious particles, which are then used to infect T-lymphocytes. The resulting cytotoxic T-lymphocytes (CTLs) are selectively activatable, and have a lessened dependence on helper T-lymphocytes (TH cells) and/or growth factors. The chimeric receptor binds to the LFA-1 ligand (intercellular adhesion molecule-1, ICAM-1 or CD54) but IC28 transmits the proliferative signal of CD28. The CTLs mediate lysis of a target cell bearing a cognate antigen, and are useful in directed cancer immunotherapy, and in treatment of viral infections

XX Sequence 3585 BP; 808 A; 981 C; 1011 G; 785 T; 0 U; 0 Other;
 SQ Query Match 68.0%; Score 20.4; DB 2; Length 3585;
 Best Local Similarity 95.5%; Pred. No. 42;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 9 CTTGTCAAGGCCATAGGCTGC 30
 Db 3578 CGGTCTGGAGGCCATAGGCTGC 3557

RESULT 21 ABV75351/C
 ID ABV75351 Standard; DNA; 3804 BP.
 XX AC ABV75351;
 XX DT 07-MAR-2003 (first entry)
 XX DE Human CD28 protein encoding DNA.
 XX DE Human CD28 coding sequence.
 XX KW Human; CD28; immune response; signaling; diabetes; autoimmune disease; immunosuppressives; dermatological; antiinflammatory; antithyroid; antirheumatic; antiarthritic; nootropic; allergy; rheumatoid arthritis; systemic lupus erythematosus; myasthenia gravis; autoimmune thyroiditis; vitiligo; alopecia; inflammatory bowel disease; Addison's disease; Graves disease; haemolytic anaemia; Sjogren's syndrome; gene; ds.
 XX OS Homo sapiens.
 XX FH Location/Qualifiers
 FT 222..884
 FT /*tag= a
 FT /product= "CD28"
 FT /transl_except= (pos:867..872,aa:Ala)
 XX PN WO20020518-A2.
 XX EN 14-NOV-2002.
 XX PD 10-MAY-2002; 2002WO-US014843.
 XX PR 10-MAY-2001; 2001US-0290097P.
 XX PN WO20026059-A2.

XX PD 29-AUG-2002.
 XX PP 15-FEB-2002; 2002WO-US004772.
 XX PR 16-FEB-2001; 2001US-0269756P.
 XX PA (GEMY) GENETICS INST INC.
 XX PI O'hara RM, Nagelin AM;
 XX DR WPI; 2002-674899/72.
 PT P-PSDB; AA019101.
 XX PS Disclosure: Page 60-61; 61PP; English.
 XX PT Downmodulating immune responses by blocking CD28-mediated signaling, useful for preventing and/or treating autoimmune disorders such as diabetes, allergic reactions, graft versus host disease, systemic lupus erythematosus.
 XX PT Disclosure: Page 60-61; 61PP; English.
 XX CC The present invention relates to a method of therapeutically downmodulating an autoimmune response or an ongoing autoimmune response, comprising administering an antigen binding portion of an anti-CD28 antibody that blocks signaling via CD28 to the subject so that an autoimmune response or an ongoing autoimmune response in the subject is downmodulated. The methods are useful in therapeutically and prophylactically downmodulating the immune response in subjects having autoimmune disorders such as diabetes, allergy and allergic reactions, transplantation rejection, graft versus host disease, systemic lupus erythematosus, rheumatoid arthritis, myasthenia gravis, autoimmune thyroiditis, vitiligo, alopecia, inflammatory bowel disease, Addison's disease, Graves disease, hemolytic anaemia and Sjogren's syndrome. The present sequence is the human CD28 coding sequence
 XX SQ Sequence 3803 BP; 1076 A; 766 C; 880 G; 1081 T; 0 U; 0 Other;
 SQ Query Match 68.0%; Score 20.4; DB 6; Length 3803;
 Best Local Similarity 95.5%; Pred. No. 43;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 CTTGTCAAGGCCATAGGCTGC 30
 Db 888 CGGTCTGGAGGCCATAGGCTGC 867

RESULT 22 ABV75351/C
 ID ABV75351 Standard; DNA; 3804 BP.
 XX AC ABV75351;
 XX DT 07-MAR-2003 (first entry)
 XX DE Human CD28 protein encoding DNA.
 XX KW T-cell; cell proliferation; cell survival; cytosatic; antiallergic; immunostimulant; immunosuppressive; gene therapy; CD28; human; gene; ds.
 XX OS Homo sapiens.
 XX Key CDS 223..885
 FT /*tag= a
 FT /product= "CD28"
 XX PN Location/Qualifiers
 XX EN 14-NOV-2002.
 XX PD 10-MAY-2002; 2002WO-US014843.
 XX PR 10-MAY-2001; 2001US-0290097P.
 XX PN WO20026059-A2.

XX (UNIW) UNIV WASHINGTON.
 PA XX Claim 1; SEQ ID NO 1509; 245pp; English.
 PI Green JM, Shaw AS;
 XX The invention relates to recombinant carcinoma associated (CA) nucleic
 DR acid sequences from mouse and human (ADA01482-ADA01094), and to
 P-PSDB; ABB82721. recombinant carcinoma associated proteins (CAP) encoded by them. The
 XX invention also encompasses expression vectors and host cells comprising a
 CA nucleic acid, a polypeptide (especially an antibody) that specifically
 binds to the protein, and a biochip comprising CA nucleic acid or
 PT fragments thereof. The sequences of the invention were identified using
 PT oncogenic retroviruses, which insert into the genome of the host organism
 PT at random. Many of these do not carry transduced host oncogenes or
 PT pathogenic trans-acting viral genes, meaning that cancer incidence is a
 PT direct consequence of the effects of proviral integration into host
 XX protooncogenes. The CA nucleic acid sequences can be used to diagnose
 PS carcinoma (especially breast cancer, prostate cancer, lymphoma or
 CC leukaemia) or a propensity to carcinoma by determination of the sequence
 CC of a CA gene, or by determination of CA gene expression in particular
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as
 CC therapeutic agents and in screening and evaluating drug candidates. The
 CC present sequence represents specifically claimed human CA nucleic acid
 CC sequence of the invention. Note: The complete sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 3804 BP; 1076 A; 766 C; 880 G; 1082 T; 0 U; 0 Other;
 SQ Query Match 68.0%; Score 20.4; DB 9; Length 3804;
 Best Local Similarity 95.5%; Pred. No. 43;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 9 CTTGTAGGAGCCATAGGCTGC 30
 Db 889 CGTGTACGGAGCCATAGGCTGC 868
 RESULT 24
 ADB72729/c
 ID ADB72729 standard; mRNA; 3804 BP.
 AC XX
 ADB72729;
 ID AC
 XX DT 04-DEC-2003 (first entry)
 XX DE Human CD28 mRNA.
 XX KW human; ss: cytosstatic; gene therapy; vaccine; carcinoma; lymphomas;
 KW cancer; neoplasm; adenocarcinoma; sarcoma.
 OS Homo sapiens.
 XX WO200308583-A2.
 XX PD 30-JAN-2003.
 XX PF 26-DEC-2001; 2001WO-US051291.
 XX PR 02-MAR-2001; 2001US-00798516.
 XX PR 23-OCT-2001; 2001US-00004113.
 XX PR 08-NOV-2001; 2001US-00052482.
 XX PR 30-NOV-2001; 2001US-00997722.
 XX PR 20-DEC-2001; 2001US-00034650.
 XX PA (SAGR-) SAGRES DISCOVERY.
 XX PI Morris DW, Engelhard BK;
 XX DR WPI; 2003-239337/23.
 XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
 PT cancers, neoplasm, adenocarcinoma, or sarcomas.
 XX Claim 1; SEQ ID NO 557; 2304pp; English.
 PS

XX The invention relates to a novel recombinant nucleic acid comprising a CC nucleotide sequence selected from any of the 660 sequences fully defined CC in the specification. A polynucleotide of the invention has cytosatic CC activity, and may have a use in gene therapy, or in a vaccine. The CC recombinant nucleic acids and polypeptides are useful for treating CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and CC carcinomas. The present sequence represents a human mRNA of the invention.

XX Sequence 3804 BP; 1076 A; 766 C; 880 G; 1082 T; 0 U; 0 Other;

Query Match 68.0%; Score 20.4%; DB 10; Length 3804;
Best Local Similarity 95.5%; Pred. No. 43;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 CTGGTCAGGAGCCATAGGTGC 30

Db 889 CGTGTCAAGGCGATAGGTGC 868

RESULT 25

ADC85471/C
ID ADC85471 Standard; DNA; 3804 BP.

XX DT 01-JAN-2004 (first entry)
XX Human Cd128 mRNA sequence.
XX Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
KW secreted; transmembrane; intracellular; ds.
XX OS Homo sapiens.
XX PN WO2003045230-A2.
XX PD 05-JUN-2003.
XX PP 02-DEC-2002; 2002WO-US0308532.
XX PR 30-NOV-2001; 2001US-00997722.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW, Engelhardt EK;
XX DR 2003-513603/48.
XX PS 1; SEQ ID NO 257; 98pp; English.

The invention relates to a recombinant nucleic acid comprising a nucleotide sequence selected from any of the 50 tables given in the specification. The carcinoma-associated (CA) genes, useful for screening for drug candidates for diagnosing or treating carcinomas. Sequences given in ADC85514 represent CA genes of the invention.

XX Sequence 3804 BP; 1076 A; 766 C; 880 G; 1082 T; 0 U; 0 Other;

Query Match 68.0%; Score 20.4%; DB 10; Length 3804;
Best Local Similarity 95.5%; Pred. No. 43;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 CTGGTCAGGAGCCATAGGTGC 30

Db 889 CGTGTCAAGGCGATAGGTGC 868

RESULT 26

ADM74586/C
ID ADM74586 standard; DNA; 3804 BP.

XX AC ADM74586;
XX DT 01-JUL-2004 (first entry)
XX DE Human carcinoma associated (CA) nucleic acid #128.

XX KW Human; carcinoma associated nucleic acid; CA nucleic acid; gene; ss;
carcinoma associated protein; CAP; carcinoma; leukaemia; lymphoma;
cytosatic.
XX OS Homo sapiens.
XX PN US2004072154-A1.
XX PD 15-APR-2004.
XX PF 30-NOV-2001; 2001US-00997722.
XX PR 22-DEC-2000; 2000US-00747377.
XX PR 02-MAR-2001; 2001US-00798586.
XX PA (MORRIS D W.
(ENGELHARD E K.
XX PI Morris DW, Engelhardt EK;
XX DR WPI; 2004-328562/30.
XX PT New carcinoma associated gene or protein, useful for preparing a
composition for diagnosing or treating carcinoma e.g., leukemia or
lymphoma.
XX PT Claim 1; SEQ ID NO 257; 29pp; English.
XX PS PS 1; SEQ ID NO 257; 29pp; English.

The invention relates to new recombinant nucleic acids. The invention also relates to a host cell comprising a recombinant nucleic acid or expression vector, an expression vector comprising recombinant nucleic acid, a recombinant protein, a method of screening for drug candidates, a method of screening for a bioactive agent capable of binding to a carcinoma associated protein (CAP) encoded by a nucleotide sequence, a method of screening for a bioactive agent capable of modulating the activity of a CAP, a method of evaluating the effect of a candidate carcinoma drug, a method of diagnosing carcinoma, a method for inhibiting the activity of a CAP, a method of treating carcinoma or neutralising the effect of a CAP and a method of diagnosing carcinoma or propensity to carcinoma. A method of evaluating the effect of candidate carcinoma drug comprises administering the drug to a patient, removing a cell sample from the patient and determining alterations in the expression or activation of a gene comprising the nucleotide sequence. A method of diagnosing carcinoma comprises determining the expression of one or more genes comprising the nucleic acid sequence in a first tissue type of a first individual and comparing the expression of the gene from a second normal tissue type from the first individual or a second unaffected individual, where a difference in the expression indicates that the first individual has carcinoma. A method of inhibiting the activity of a CAP comprises binding an inhibitor to the CAP. Treating carcinoma comprises administering to a patient an inhibitor of CAP. Neutralising the effect of CAP comprises contacting an agent specific for the CAP. The polypeptide specifically binds to the protein encoded by the nucleic acid. It comprises an antibody that specifically binds to the protein encoded by the nucleic acid, the nucleic acids are useful for preparing a composition for diagnosing or treating carcinoma e.g., leukaemia or lymphoma. This sequence represents a human carcinoma associated (CA) nucleic acid of the invention. Note: The sequence data obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

SQ Sequence 3804 BP; 1076 A; 766 C; 880 G; 1082 T; 0 U; 0 Other;
 Query Match 68.0%; Score 20.4; DB 12; Length 3804;
 Best Local Similarity 95.5%; Pred. No. 43;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 CTTGTCAGGAGCGATAGGTGC 30
 Db 889 CGTGTCAAGAGCGATAGGTGC 868

RESULT 27
 ID ADN05855/c
 ADN05855 standard; cDNA; 3804 BP.
 XX AC ADN05855;
 XX DT 01-JUL-2004 (first entry)
 XX DE Antipsoriatic cDNA sequence #1159.
 XX KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
 OS Homo sapiens.
 XX PN WO2004028479-A2.
 XX WPI; 2004-028479-A2.
 XX PD 08-APR-2004.
 XX PF 25-SEP-2003; 2003WO-US030907.
 XX PR 25-SEP-2002; 2002US-0414006P.
 XX PA (GETH) GENENTECH INC.

XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
 PI Wu TD;
 XX WPI; 2004-305105/28.
 DR P-PSDB; ADN058556.

XX PT New PRO nucleic acid or polypeptide, useful for preparing a composition
 PT pharmaceutical composition for diagnosing or treating psoriasis in a
 PT mammal.
 XX PS Claim 1; SEQ ID NO 2250; 3069pp; English.

CC The invention relates to novel polynucleotide and polypeptides for
 CC treating psoriasis or a sequence having at least 80% identity to the
 CC above sequences. The nucleic acid is useful for preparing a composition
 CC for diagnosing or treating psoriasis in a mammal. This sequence
 CC corresponds to one of the polynucleotides of the invention.

XX SQ Sequence 3804 BP; 1076 A; 766 C; 880 G; 1082 T; 0 U; 0 Other;
 Query Match 68.0%; Score 20.4; DB 12; Length 3804;
 Best Local Similarity 95.5%; Pred. No. 43;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 CTTGTCAGGAGCGATAGGTGC 30
 Db 889 CGTGTCAAGAGCGATAGGTGC 868

RESULT 28
 ADN021976/c
 ID ADN021976 standard; DNA; 3806 BP.
 XX AC AAD21976;
 XX DT 12-FEB-2002 (first entry)
 XX DE Human CD28 DNA.

XX XX Human; cytotoxic T lymphocyte antigen-4; CTLA-4; CD28; AIDS; therapy;
 KW Acquired Immune Deficiency Syndrome; neuroprotective; dermatological;
 KW immune response; organ transplantation; autoimmune disease; allergy; SLE;
 KW systemic lupus erythematosus; multiple sclerosis; tumour vaccination;
 KW immunodeficiency disease; DiGeorge Syndrome; cancer; ds.
 XX Homo sapiens.
 XX XX Location/Qualifiers
 FH Key
 FT CDS
 FT 223 . 885
 FT /*tag= a
 /product= "Human CD28 protein"

XX XX PN WO200179300-A1.
 XX XX PD 25-OCT-2001.
 XX PP 12-APR-2001; 2001IWO-US012275.
 XX PR 12-APR-2000; 2000US-0196851P.
 XX PA (GEMY) GENETICS INST INC.
 XX PI Bluestone JA, Collins M, Whitters M, Griffin M, Kranz D;
 XX DR WPI; 2002-017603/02.
 DR P-PSDB; AAE13721.

XX XX PT New construct for downmodulating immune response in a subject, has
 PT exposed surface attached with antigen-binding portion of antibody that
 PT binds to cytotoxic T lymphocyte antigen-4 and major histocompatibility
 PT molecule.

XX XX PS Disclosure; Page 93-94; 98pp; English.

XX XX CC The invention relates to a construct for downmodulating immune response
 CC in a subject. The construct comprises an exposed surface attached with an
 CC antigen-binding portion of an antibody that binds to a cytotoxic T
 CC lymphocytic antigen (CTLA)-4 or CD28 expressed on T-cell of the subject,
 CC and a major histocompatibility complex (MHC) molecule. The construct is
 CC useful for treating an individual afflicted with a disease or disorder
 CC that would benefit from downregulation of immune response. Downmodulation
 CC of the immune response is useful to downmodulate the immune response in
 CC situations of tissue, skin and organ transplantation, graft-versus-host
 CC disease or in autoimmune diseases such as systemic lupus erythematosus
 CC and multiple sclerosis. The construct is useful for inhibiting immune
 CC cell activation and prevents production of autoantibodies or cytokines
 CC which may be involved in disease process. Inhibition of immune cell
 CC activation is useful in the treatment of allergy and allergic reactions
 CC e.g., by inhibiting IgE production. The construct is also useful for
 CC enhancing desirable immune response in a subject e.g., in situations of
 CC tumour vaccination, in viral immunity or in immunodeficiency diseases
 CC such as AIDS and DiGeorge Syndrome. Induction/enhancement of immune cell
 CC construction results in increased tumour destruction in cancer patients. The
 CC present construct is also useful in treating infectious diseases. The present
 CC sequence is human CD28 DNA

XX XX SQ Sequence 3806 BP; 1079 A; 765 C; 881 G; 1081 T; 0 U; 0 Other;
 Query Match 68.0%; Score 20.4; DB 6; Length 3806;
 Best Local Similarity 95.5%; Pred. No. 43;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 CTTGTCAGGAGCGATAGGTGC 30
 Db 889 CGTGTCAAGAGCGATAGGTGC 868

RESULT 29
 ADA02990/C
 ID ADA02990 standard; DNA; 51365 BP.

AC ADA02990;
 XX 06-NOV-2003 (first entry)
 XX Human CD28 carcinoma associated gene, SEQ ID NO:1508.
 DE WO2003057146-A2.
 KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
 KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
 KW gene; ds.
 XX Homo sapiens.
 XX OS
 XX PN
 XX WO2003057146-A2.
 PD 17-JUL-2003.
 XX 26-DEC-2002; 2002WO-US041414.
 XX 26-DEC-2001; 2001US-00035832.
 PR
 PA (SAGR-) SAGRES DISCOVERY.
 XX Morris DW,
 PI Engelhard BK.
 XX DR
 XX WPI; 2003-239337/23.
 PT New recombinant nucleic acid encoding carcinoma associated protein,
 PT useful for preparing compositions for treating carcinomas.
 XX PS
 XX SEQ ID NO 1508; 245pp; English.
 XX PT
 XX DR
 XX WPI; 2003-587068/55.
 PT New recombinant nucleic acid encoding carcinoma associated protein,
 PT useful for preparing compositions for treating carcinomas.
 XX PS
 XX SEQ ID NO 1508; 245pp; English.
 CC The invention relates to recombinant carcinoma associated (CA) nucleic
 CC acid sequences from mouse and human (ADA01482; ADA03094), and to
 CC recombinant carcinoma associated proteins (CAP), encoded by them. The
 CC invention also encompasses expression vectors and host cells comprising a
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
 CC binds to the protein, and a biochip comprising CA nucleic acid or
 CC fragments thereof. The sequences of the invention were identified using
 CC oncogenic retroviruses, which insert into the genome of the host organism
 CC at random. Many of these do not carry transduced host oncogenes or
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
 CC direct consequence of the effects of proviral integration into host
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
 CC leukaemia) or a propensity to carcinoma by determination of the sequence
 CC of a CA gene, or by determination of CA gene expression in particular
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as
 CC therapeutic agents and in screening and evaluating drug candidates. The
 CC present sequence represents a specifically claimed human CA nucleic acid
 CC sequence of the invention. Note: The complete sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 51365 BP; 15176 A; 10059 C; 10474 G; 15656 T; 0 U; 0 Other;
 XX Query Match 68.0%; Score 20.4; DB 9; Length 51365;
 XX Best Local Similarity 95.5%; Pred. No. 62;
 XX Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX Qy 9 CTTGTCAAGGAGCATAGGGTGC 30
 XX Db 38448 CGTGTCAAGGAGCATAGGTG 38427
 RESULT 31
 ADC85470/C
 ID ADC85470 standard; DNA; 51365 BP.
 ID ADC85470;
 AC AC
 AC XX
 AC DT 01-JAN-2004 (first entry)
 AC XX
 AC DE Human Cd28 genomic sequence.
 AC XX
 AC KW Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
 AC XX
 AC OS Homo sapiens.
 AC XX
 AC PN WO2003045230-A2.
 AC XX
 AC PD 05-JUN-2003.
 AC XX
 AC PP 02-DEC-2002; 2002WO-US038582.
 AC XX
 AC PR 30-NOV-2001; 2001US-00997722.
 AC XX
 AC PA (SAGR-) SAGRES DISCOVERY.

XX Morris DW, Engelhardt EK;
 XX WPI; 2003-513603/48.

XX New recombinant nucleic acid comprising a nucleotide sequence of any of
 PT the carcinoma-associated (CA) genes, useful for screening for drug
 PT candidates for diagnosing or treating carcinomas.
 XX

PS Claim 1; SEQ ID NO 256; 983pp; English.

XX The invention relates to a recombinant nucleic acid comprising a
 CC nucleotide sequence selected from any of the fully defined carcinoma-
 CC associated (CA) genes from the 50 tables given in the specification. The
 CC proteins are secreted, transmembrane or intracellular proteins. The
 CC recombinant nucleic acids are useful for screening for drug candidates
 CC for diagnosing or treating carcinomas. Sequences given in ADC85215-
 CC ADC85514 represent CA genes of the invention.
 XX Sequence 51365 BP; 15176 A; 10059 C; 10474 G; 15656 T; 0 U; 0 Other;

Query Match 68.4%; Score 20.4; DB 10; Length 51365;
 Best Local Similarity 95.5%; Pred. No. 62;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 CTTGCTCAGGAGCGATAGGCCTGC 30
 Db 38448 CGTGTCAAGGCGATAGGCCTGC 38427

RESULT 32
 ADM74585 C
 ID ADM74585 standard; DNA; 51365 BP.
 XX AC ADM74585;
 XX DT 01-JUL-2004 (first entry)
 DE Human carcinoma associated (CA) nucleic acid #127.
 XX Human; carcinoma associated nucleic acid; CA nucleic acid; gene; ds;
 KW carcinoma associated protein; CAP; carcinoma; leukaemia; lymphoma;
 KW cytostatic.
 XX OS Homo sapiens.
 XX US2004072154-A1.
 XX 15-APR-2004.
 XX 30-NOV-2001; 2001US-00997722.
 PR 22-DEC-2000; 2000US-00747377.
 PR 02-MAR-2001; 2001US-00798586.
 PA (MORRIS D W.
 PA (ENGE/) ENGELHARD E K.
 XX Morris DW, Engelhardt EK;
 XX WPI; 2004-328562/30.

XX New carcinoma associated gene or protein, useful for preparing a
 PT composition for diagnosing or treating carcinoma e.g., leukemia or
 PT lymphoma.
 XX Claim 1; SEQ ID NO 256; 29pp; English.

XX The invention relates to new recombinant nucleic acids. The invention
 CC also relates to a host cell comprising a recombinant nucleic acid or
 CC expression vector, an expression vector comprising a recombinant nucleic
 CC acid, a recombinant protein, a method of screening for drug candidates,
 CC method of screening for a bioactive agent capable of binding to a

CC carcinoma associated protein (CAP) encoded by a nucleotide sequence, a
 CC method of screening for a bioactive agent capable of modulating the
 CC activity of a CAP, a method of evaluating the effect of a candidate
 carcinoma drug, a method of diagnosing carcinoma, a method of inhibiting
 the activity of a CAP, a method of treating carcinoma, a method of
 neutralising the effect of a CAP and a method of diagnosing carcinoma or
 propensity to carcinoma. A method of evaluating the effect of a candidate
 carcinoma drug comprises administering the drug to a patient, removing a
 CC cell sample from the patient and determining alterations in the
 CC expression or activation of a gene comprising the nucleotide sequence. A
 CC method of diagnosing carcinoma comprises determining the expression of
 CC one or more genes comprising the nucleic acid sequence in a first tissue
 CC type of a first individual and comparing the expression of the gene from
 CC a second normal tissue type from the first individual or a second
 CC unaffected individual, where a difference in the expression indicates
 CC that the first individual has carcinoma. A method of inhibiting the
 CC activity of a CAP comprises binding an inhibitor to the CAP. Treating
 CC carcinomas comprises administering to a patient an inhibitor of CAP.
 CC Neutralising the effect of a CAP comprises contacting an agent specific
 CC for the CAP. The polypeptide specifically binds to the protein encoded by
 CC the nucleic acid. It comprises an antibody that specifically binds to the
 CC protein encoded by the nucleic acid. The nucleic acids are useful for
 CC preparing a composition for diagnosing or treating carcinoma e.g.,
 CC leukaemia or lymphoma. This sequence represents a human carcinoma
 CC associated (CA) nucleic acid of the invention. Note: The sequence data
 CC for this patent did not form part of the printed specification but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html.

SQ Sequence 51365 BP; 15176 A; 10059 C; 10474 G; 15656 T; 0 U; 0 Other;

Query Match 68.0%; Score 20.4; DB 12; Length 51365;
 Best Local Similarity 95.5%; Pred. No. 62;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 CTTGCTCAGGAGCGATAGGCCTGC 30
 Db 38448 CGTGTCAAGGCGATAGGCCTGC 38427

RESULT 33
 ADL67233
 ID ADL67233 standard; DNA; 36 BP.
 XX AC ADL67233;
 XX DT 20-MAY-2004 (first entry)
 DE Human CD28 intracellular domain amplifying primer #2.
 XX KW T cell receptor; TCR; CD3 zeta chain; co-stimulatory signalling region;
 KW binding element; immunostimulant; therapy; cancer; primer; ss.
 XX OS Homo sapiens.
 XX PN US2004043401-A1.
 XX PD 04-MAR-2004.
 XX PF 28-MAY-2003; 2003US-00448256.
 XX PR 28-MAY-2002; 2002US-0383872P.
 XX (SLOK) SLOAN KETTERING INST CANCER RES.
 XX PI Sadelain M, Brentjens R, Maher J;
 XX DR WPI; 2004-225696/21.
 XX New nucleic acid polymer encoding a chimeric T cell receptor having a
 PT zeta chain portion, useful for treating disorders where the immune
 PT response needs to be induced, such as cancer.

PS Example 1; SEQ ID NO 8; 25pp; English.
 XX The invention relates to a nucleic acid polymer encoding a chimeric T
 CC cell receptor (TCR) which comprises human CD3 zeta chain intracellular
 CC domain, co-stimulatory signalling region and binding
 CC specifically with a selected target. The methods and
 CC compositions of the invention are useful for treating disorders where the
 CC immune response needs to be induced, such as cancer. The present sequence
 CC is a PCR primer used to amplify human CD28 costimulatory molecule
 CC intracellular domain.
 XX Sequence 36 BP; 7 A; 9 C; 14 G; 6 T; 0 U; 0 Other;
 SQ Query Match 67.3%; Score 20.2; DB 12; Length 36;
 Best Local Similarity 88.0%; Prd. No. 27;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 AC GATCTTGTAGGAGCGATAAGGCTGC 30
 DB 2 GTCGAGTAGGAGGCGATAAGGCTGC 26

RESULT 34
 ABN88948 ID ABN88948 standard; DNA; 31 BP.
 XX AC ABN88948;
 XX DT 22-AUG-2002 (first entry)
 XX DE Human CD28 35S A antisense primer.
 XX KW Monoclonal antibody 8H9; tumour-associated antigen; 8H9 antigen;
 KW immunotherapy; antitumour; antibody-based therapy; tumour; cytotoxic;
 KW tumour bearing potential; cancer; Ewing's sarcoma; primer; 8S.
 XX OS Homo sapiens.
 XX OS Synthetic.
 PN WO200232375-A2.
 XX PD 25-APR-2002.
 XX PF 18-OCT-2001; 2001WO-US032555.
 XX PR 18-OCT-2000; 2000US-0241344P.
 XX PR 17-OCT-2001; 2001US-0330368P.
 PA (SLOK) SLOAN KETTERING INST CANCER RES.
 XX PS Cheung NV;
 XX PI Page 142; 209pp; English.
 DR WPI; 2002-479645/51.

PT Novel tumor-associated antigen recognized by murine monoclonal antibody 8H9, expressed on cell membranes of broad spectrum of tumors, useful for producing antibodies capable of inhibiting tumor cell growth.
 XX Example: Page 142; 209pp; English.
 XX The present invention describes a protein (I), in particular tumour-associated antigen 58 kilo Dalton (kDa) in molecular weight, reacting specifically with monoclonal antibody (MAB) 8H9. Also described: (1) a composition (C) comprising MAB 8H9 or its derivative; (2) an antibody produced by immunising (I) or its specific portion; and (3) an antibody (II) produced by using (I). (I) has antitumour activity and can be used in antibody-based therapy. (I) can be used for producing a Mab, and to evaluate the tumour bearing potential of a subject, by measuring the expression of 8H9 antigen in the subject, where an increased expression of the antigen indicates higher tumour bearing potential of the subject. (II) or (C) comprising 8H9 or its derivative are useful for inhibiting the growth of tumour cells *in vitro* or *in vivo*, and the antibody indirectly coupled to a cytotoxic agent (radioisotope) is useful for

CC reducing tumour cells in a subject. The present sequence represents a
 CC primer which is used in the exemplification of the present invention.
 XX Sequence 31 BP; 5 A; 6 C; 12 G; 7 T; 0 U; 1 Other;
 SQ Query Match 66.7%; Score 20; DB 6; Length 31;
 Best Local Similarity 100%; Pred. No. 32;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 11 TGTCAAGGAGCGATAAGGCTGC 30
 Db 12 TGTCAAGGAGCGATAAGGCTGC 31

RESULT 35
 ACF80096 ID ACF80096 standard; DNA; 31 BP.
 XX AC ACF80096;
 XX DT 15-JAN-2004 (first entry)
 XX DE Human CD8a leader 355A antisense PCR primer.
 XX KW Monoclonal antibody; 8H9; antitumour; cytostatic; gene therapy; human;
 KW CD28; PCR; primer; 8S.
 XX OS Homo sapiens.
 XX PN WO2003075846-A2.
 XX PD 18-SEP-2003.
 XX PF 06-MAR-2003; 2003WO-US007004.
 XX PR 08-MAR-2002; 2002US-00097558.
 PR 17-OCT-2002; 2002US-US0273762.
 PR 17-OCT-2002; 2002WO-US0333331.
 XX PA (SLOK) SLOAN KETTERING INST CANCER RES.
 XX PI Cheung NV;
 XX DR WPI; 2003-731784/69.

XX Composition comprising the monoclonal antibody 8H9 useful for reducing tumor cells in a subject, evaluating the tumor bearing potential of a subject or screening new anti-tumor compounds.
 XX Disclosure; Page 124; 190pp; English.
 XX The present sequence is that of 355A antisense primer for human CD28, incorporating an XbaI restriction site. It was used with primer 355S sense primer (see ACF80095) in the PCR amplification of an hCD8a leader-BH9 scFv-CD8 construct. The construct was used in a series of experiments in which an anti-idiotype antibody was used to enhance scFv chimeric immune receptor gene transduction and clonal expansion of human lymphocytes. The invention provides a composition comprising 8H9 or its derivative, such as an scFv or scFv-FC, and a suitable carrier. The antibody may comprise the CDRs of 8H9 with its remaining sequence being human. Also claimed are nucleic acids encoding the antibodies, vectors, cells comprising the vectors, and methods of producing the antibodies. A method for directly killing or delivering a drug, DNA or RNA to cells bearing the antigen recognised by 8H9 or to image cells or tumours bearing this antigen using an 8H9 scFv or a 8H9-scFv-modified cell or liposome is also claimed. The antibody is used in claimed methods of inhibiting the growth of tumour cells, imaging a tumour, evaluating the tumour bearing potential of a subject, or (when coupled to a cytotoxic agent) reducing tumour cells

XX Sequence 31 BP; 5 A; 6 C; 12 G; 7 T; 0 U; 1 Other;
 SQ Query Match 66.7%; Score 20; DB 10; Length 31;

Best Local Similarity 100.0% ; Pred. No. 32; Matches 20; Conservative 0; Mismatches 0;	Indels 0;	Gaps 0;	KW	Cell activation; chimeric receptor; DNA delivery; CD28; human; cancer; therapy; PCR; primer; ss.
Qy 11 TGTCTGGACCGATAGGCTGC 30			XX	Synthetic.
Db 12 TGTCTGGACCGATAGGCTGC 31.			OS	Homo sapiens.
			XX	WO9723613-A2.
			PN	
RESULT 36			XX	
ACC70422			PD	03-JUL-1997.
ID ACC70422 standard; DNA; 31 BP.			XX	
XX			PF	23-DEC-1996;
AC ACC70422;			XX	96WO-GB003209.
XX			PR	21-DEC-1995;
DT 11-AUG-2003 (first entry)			XX	95GB-00026131.
DE PCR primer used to amplify human CD28 DNA.			PA	(CLLT) CELLTECH THERAPEUTICS LTD.
XX			XX	
KW Single chain antibody; scFv; antigen; CD8a; CD28; PCR; primer; ss.			PI	Babbington CR, Lawson ADG, Weir ANC, Finney HM,
XX			XX	WPI; 1997-351052/32.
OS Synthetic.			DR	
XX			XX	New DNA systems for activating cells - comprising DNA coding for a chimeric receptor comprising 2 or more different cytoplasmic signalling components.
PN WO2003033670-A2.			PS	Example 1; Fig 3; 90pp; English.
XX			XX	Primer P3241 was used with primer P3240 (see AAT90525) in the PCR cloning
PD 24-APR-2003.			CC	of the human CD28 transmembrane and intracellular components using human
XX			CC	leukocyte cDNA as template. P3241 introduces a 3' EcoRI site and P3240
PP 17-OCT-2002; 2002WO-US0333331.			CC	introduces a 5' BamHI site. The PCR product was subcloned into
XX			CC	pBluescript SK+ to form a CD28 cassette. Primer P3241 was also used with
PR 17-OCT-2001; 2001IUS-0330396P.			CC	primer S0146 (see AAT90533) to produce a hinge-CD28 cassette. These
PR 18-OCT-2001; 2001MO-US031565.			CC	cassettes can be utilised in novel chimeric genes (see AAT90510-11 and
PR 08-MAR-2002; 2002IUS-0009755B.			CC	AAW26647-48 and AAW26650- 51) encoding recombinant chimeric receptors (see AAT90513-14) useful in cell activation processes, e.g. for the treatment
XX			CC	of cancer
PA (SLOK) SLOAN KETTERING INST CANCER RES.			SQ	Sequence 30 BP; 9 A; 5 C; 9 G; 7 T; 0 U; 0 Other;
PA Cheung NV, Guo H;			Query Match	Score 64.7%; Score 19.4; DB 2; Length 30;
PI XX			Best Local Similarity	95.2%; Pred. No. 60;
DR 2003-393522/37.			Matches	Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX			Qy	10 TTTCTAGAGGCTAGGCTGC 30
PT Preparing single chain antibody directed against an antigen, by selecting clones using anti-idiotype antibody directed to antibody specific for the antigen.			Db	7 TTCTAGAGGCTAGGCTGC 27
PT Disclosure; Page 46; 136pp; English.				
XX				RESULT 38
CC The specification describes a method of making a single chain antibody CC directed against an antigen, where the selection of clones is made based upon interaction of those clones with an appropriate anti-idiotype. CC The method is useful for making a single chain antibody directed against CC an antigen. PCR primers ACC70421-22 were used to construct a chimeric CC immune receptor, comprising the human CD8a leader, a scFv and CD28 CC domains		AAT90529/C		
XX			ID	AAT90529 standard; DNA; 64 BP.
SQ Sequence 31 BP; 5 A; 6 C; 12 G; 7 T; 0 U; 1 Other;			XX	
Query Match			AC	AAT90529;
Best Local Similarity 66.7% ; Score 20; DB 10; Length 31;			XX	
Mismatches 20; Conservative 0; Mismatches 0;			DT	12-FEB-1998 (first entry)
Indels 0;			XX	T cell receptor zeta chain-CD28 fusion cassette primer P3303.
Gaps 0;			DE	KW Cell activation; chimeric receptor; DNA delivery; CD28; T cell receptor;
			XX	KW human; cancer; therapy; PCR; primer; ss.
			OS	Synthetic.
			OS	Homo sapiens.
XX			XX	WO9723613-A2.
XX			PN	
AC AAT90526;			XX	
XX			PD	03-JUL-1997.
DT 12-FEB-1998 (first entry)			XX	
XX			PR	23-DEC-1996;
DE Human CD28 PCR primer P3241.			XX	96WO-GB003209.
XX			PA	95GB-00026131.
				(CLLT) CELLTECH THERAPEUTICS LTD.

XX Bebbington CR, Lawson ADG, Weir ANC, Finney HM;
 PI WPI; 197-351052/32.

XX New DNA systems for activating cells - comprising DNA coding for a
 PT chimeric receptor comprising 2 or more different cytoplasmic signalling
 PT components.
 XX Example 1: Fig 3; 90pp; English.

XX Primers P3301, P3302, P3303, P3304, P3305 and P3306 (see AT90527-32)
 CC were used for the PCR assembly of a human T cell receptor zeta-CD28
 CC fusion cassette. The 3' end of zeta, starting at the native Sty1 site,
 CC and the intracellular component of human CD28 were PCR assembled such
 CC that the zeta stop codon was removed and an in-frame fusion protein would
 CC be translated. The PCR product was subcloned into pBluescript SK+
 CC containing a CTRM01 TCR zeta chimeric receptor construct, replacing the 3'
 CC end of zeta. The zeta-CD28 fusion cassette can be utilised in novel
 CC chimeric genes (see AT90510 and AT90513) encoding recombinant chimeric
 CC receptors (see AAW6647 and AAW6650) useful in cell activation
 CC processes, e.g. for the treatment of cancer
 XX Sequence 64 BP; 16 A; 25 C; 10 G; 13 T; 0 U; 0 Other;
 SQ

Query Match 64.7%; Score 19.4; DB 2; Length 64;
 Best Local Similarity 95.2%; Pred. No. 67;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 10 TTGTCAAGGGCATAGGCTGC 30
 Db 58 TTCTCAGGGCGATAGGCTGC 38

RESULT 39
 ADA69416 ID ADA69416 standard; DNA; 852 BP.
 XX AC ADA69416;
 XX DT 20-NOV-2003 (first entry)
 XX Rice gene, SEQ ID 2739.
 XX Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.
 XX Oryza sativa.
 OS WO200300898-A1.
 PN 03-JAN-2003.
 XX 22-JUN-2001; 2001WO-IB001105.
 XX 22-JUN-2001; 2001WO-IB001105.
 XX (SIGN) SYNTENTA PARTICIPATIONS AG.
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 DR WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX Claim 6; SEQ ID NO 2739; 89pp; English.

XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 XX comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX SQ Sequence 852 BP; 219 A; 230 C; 267 G; 136 T; 0 U; 0 Other;
 Query Match 64.7%; Score 19.4; DB 8; Length 852;
 Best Local Similarity 79.3%; Pred. No. 98;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 CGAGGATCTGAGGAGGATGGCTGC 30
 Db 417 CCAGGAATCTTCAGGAGGAGGCTGC 445

RESULT 40
 AAS81566 ID AAS81566 standard; cDNA; 1156 BP.
 XX AC AAS81566;
 XX DT 13-FEB-2002 (first entry)
 DE DNA encoding novel human diagnostic protein #17370.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; RS.
 XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PP 30-MAR-2001; 2001WO-US0008631.
 XX PR 31-MAR-2000; 2000US-00540217.
 XX DR 23-AUG-2000; 2000US-00649167.
 XX (HYSEQ INC.) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 PI XX DR WPI; 2001-639362/73.
 DR P-PSDB; ABG17379.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 1; SEQ ID NO 17370; 103PP; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity

Search completed: January 7, 2005, 11:40:49
Job time : 230.421 secs

Result No.	Score	Query	Match	Length	DB	ID	Description
C 1	20.4	68.0	875	3	US-08-549-846-1		Sequence 1, Appli
C 2	20.4	68.0	1064	2	US-08-529-8788-40		Sequence 40, Appli
C 3	20.4	68.0	1514	4	US-09-023-1655-1316		Sequence 1316, Appli
C 4	20.4	68.0	2405	3	US-08-549-846-3		Sequence 3, Appli
C 5	20.4	68.0	3585	3	US-08-549-846-2		Sequence 2, Appli
C 6	1.9	63.3	20	2	US-08-529-8788-18		Sequence 18, Appli
C 7	1.9	63.3	1189	4	US-09-716-129-46		Sequence 46, Appli
C 8	1.9	63.3	4758	4	US-09-774-528-129		Sequence 129, Appli
C 9	1.8	60.0	29	3	US-08-112-946A-7		Sequence 7, Appli
C 10	1.8	60.0	268	4	US-09-270-767-1027		Sequence 1027, Appli
C 11	1.8	60.0	268	4	US-09-270-767-16309		Sequence 16309, Appli
C 12	1.8	60.0	746	4	US-09-270-767-2215		Sequence 2215, Appli
C 13	1.8	60.0	746	4	US-09-270-767-17497		Sequence 17497, Appli
C 14	1.8	60.0	1498	4	US-09-270-767-15260		Sequence 15260, Appli
C 15	17.8	59.3	774	4	US-09-181-339-4		Sequence 4, Appli
C 16	17.8	59.3	1974	4	US-09-221-017B-726		Sequence 726, Appli
C 17	17.4	58.0	1758	4	US-09-252-991A-7065		Sequence 7065, Appli
C 18	17.4	58.0	2910	4	US-09-252-991A-7009		Sequence 7009, Appli
C 19	17.4	58.0	3186	4	US-09-252-991A-7120		Sequence 7120, Appli
C 20	17.2	57.3	351	4	US-09-513-999C-25301		Sequence 25301, Appli
C 21	17.2	57.3	534	4	US-09-270-767-4139		Sequence 4139, Appli
C 22	17.2	57.3	534	4	US-09-270-767-19421		Sequence 19421, Appli
C 23	17.2	57.3	939	4	US-09-489-039-3420		Sequence 3420, Appli
C 24	17.2	57.3	1492	4	US-09-741-238-11		Sequence 11, Appli
C 25	17.2	57.3	1692	4	US-09-600-991-3		Sequence 3, Appli
C 26	17.2	57.3	1709	4	US-09-600-991-21		Sequence 21, Appli
C 27	17.2	57.3	2021	1	US-07-882-925A-1		Sequence 3, Appli

RESULT 2
US -08-529-878B-40/c
Patent No. 5932556
GENERAL INFORMATION:
APPLICANT: Tam, Robert C.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
NUMBER OF SEQUENCES: 48
CORESPONDENCE ADDRESS:
ADDRESSEE: Crockett & Fish
CITY: La Habra
STATE: California
ZIP: 90631

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HEREWITH
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeiller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1316:
SEQUENCE CHARACTERISTICS:
LENGTH: 1514 base Pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 9338444
US-09-023-655-1316

INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 1064 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..309
US-08-529-878B-40

Query Match 68.0%; Score 20.4; DB 2; Length 1064;
Best Local Similarity 95.5%; Pred. No 2.7;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
9 CTGTCAAGGAGCATAGGCTGC 30
316 CGTGTCAAGGAGCATAGGCTGC 295

RESULT 3
US 09-023-655-1316/c
Sequence 1316, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart G.
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCITE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: 01-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: DYLAN, TYLER
REGISTRATION NUMBER: 37,612
REFERENCE/DOCKET NUMBER: 22627-20013.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0732

Query Match 68.0%; Score 20.4; DB 4;
Best Local Similarity 95.5%; Pred. No. 2.9;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 9 CTGTCAAGGAGCATAGGCTGC 30
Db 766 CGTGTCAAGGAGCATAGGCTGC 745

RESULT 4
US-08-549-846-3/c
Sequence 3, Application US/08549846
Patent No. 6083751
GENERAL INFORMATION:
APPLICANT: FELDHAUS, ANDREW L.
APPLICANT: JONES, LORI A.
TITLE OF INVENTION: CHIMERIC RECEPTORS FOR THE GENERATION OF
TITLE OF INVENTION: SELECTIVITY-ACTIVATABLE T(h)-INDEPENDENT CYTOTOXIC T CELLS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94104-1018

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/549,846
FILING DATE: 01-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: DYLAN, TYLER
REGISTRATION NUMBER: 37,612
REFERENCE/DOCKET NUMBER: 22627-20013.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0732

TELEX: 706141 MRSNFOERS SFO
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2405 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-549-846-3

Query Match Similarity 68.0%; Score 20.4; DB 3; Length 2405;
 Best Local Similarity 95.5%; Pred. No. 3.2;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 CTTGTCAGGAGCGATAGGGTGC 30
 Db 2398 CGTGTCAAGGAGCTAGGGTGC 2377

RESULT 5
 US-08-549-846-2/C
 Sequence 2, Application US/08549846
 Patent No. 6083751

GENERAL INFORMATION:
 APPLICANT: FELDHAUS, ANDREW L.
 APPLICANT: JONES, LORI A.
 TITLE OF INVENTION: CHIMERIC RECEPTORS FOR THE GENERATION OF
 TITLE OF INVENTION: SELECTIVITY ACTIVATABLE T(h)-INDEPENDENT CYTOTOXIC T CELLS
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FORSTER
 STREET: 755 PAGE MILL ROAD
 CITY: PALO ALTO
 STATE: CA
 COUNTRY: USA
 ZIP: 94304-1018

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patient In Release #1.0, Version #1.3.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/549, 846
 FILING DATE: 01-NOV-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: DYLAN, TILER
 REGISTRATION NUMBER: 37, 612
 REFERENCE/DOCKET NUMBER: 22627-20013.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 494-0792
 TELEX: 706141 MRSNFOERS SFO
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3585 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-549-846-2

Query Match Similarity 68.0%; Score 20.4; DB 3; Length 3585;
 Best Local Similarity 95.5%; Pred. No. 3.6;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 CTTGTCAGGAGCGATAGGGTGC 30
 Db 3578 CGTGTCAAGGAGCTAGGGTGC 3557

RESULT 6
 US-08-549-870B-18
 Sequence 18, Application US/08529870B
 Patent No. 5932556

GENERAL INFORMATION:
 APPLICANT: Tam, Robert C.
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
 REGULATION OF CD28 EXPRESSION
 NUMBER OF SEQUENCES: 48
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Crockett & Fish
 STREET: 3000 S. Augusta Court
 CITY: La Habra
 STATE: California
 COUNTRY: United States of America
 ZIP: 90631

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WordPerfect 6.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/529, 878B
 FILING DATE: 13-SEP-1995
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Fish, Robert D.
 REGISTRATION NUMBER: 33, 880
 FILE NUMBER: 213/003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 714-525-3433
 TELEX: 714-525-3303
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 US-08-529-878B-18

Query Match Similarity 63.3%; Score 19; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.6;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 TGTCAAGGAGCGATAGGTG 29
 Db 2 TGTCAAGGAGCGATAGGTG 20

RESULT 7
 US-09-716-129-46/C
 Sequence 46, Application US/09716129
 Patent No. 6632920

GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: 36 Human Secreted Proteins
 FILE REFERENCE: P2025PI
 CURRENT APPLICATION NUMBER: US/09/716,129
 CURRENT FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/076,053
 PRIOR FILING DATE: 1998-02-26
 PRIOR APPLICATION NUMBER: 60/076,057
 PRIOR FILING DATE: 1998-02-26
 PRIOR APPLICATION NUMBER: 60/076,052
 PRIOR FILING DATE: 1998-02-26
 PRIOR APPLICATION NUMBER: 60/076,054
 PRIOR FILING DATE: 1998-02-26
 PRIOR APPLICATION NUMBER: 60/076,051
 PRIOR FILING DATE: 1998-02-26
 NUMBER OF SEQ ID NOS: 186
 SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 46
 LENGTH: 1189
 TYPE: DNA
 ORGANISM: Homo sapiens

US-09-716-129-46
 Query Match 63.3%; Score 19; DB 4; Length 1189;
 Best Local Similarity 81.5%; Pred. No. 13;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 3 GAGGATCTTCAGGAGCGTAGGGCTG 29
 Db 361 GAGGATTTCAGGAGCGTAGGGCTG 335

RESULT 8
 US-09-774-528-129/C
 Sequence 129, Application US/09774528
 Patent No. 6743619

GENERAL INFORMATION:
 APPLICANT: Tang, Y, Tom
 APPLICANT: Zhou, Ping
 APPLICANT: Goodrich, Ryle
 APPLICANT: Liu, Changhua
 APPLICANT: Asundi, Vinod
 APPLICANT: Ren, Feiyan
 APPLICANT: Zhang, Jie
 APPLICANT: Zhao, Qing A.
 APPLICANT: Yang, Yonghong
 APPLICANT: Xue, Aidong J.
 APPLICANT: Wenzman, Tom
 APPLICANT: Wang, Jian-Rui
 APPLICANT: Wang, Dunrui
 APPLICANT: Drmanac, Radivoje T.

TITLE OF INVENTION: No. 674319el Nucleic Acids and
 TITLE OF INVENTION: Polypeptides

FILE REFERENCE: 802
 CURRENT FILING DATE: 2001-01-30
 NUMBER OF SEQ ID NOS: 441
 SEQ ID NO 129
 SOFTWARE: pt_FL_genes Version 2.0
 LENGTH: 4758
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (167) .. (4381)
 US-09-774-528-129

Query Match 63.3%; Score 19; DB 4; Length 4758;
 Best Local Similarity 81.5%; Pred. No. 18;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 4 AGGATCTTCAGGAGCGTAGGGCTG 30
 Db 897 AGGATCTCCCTCAGGAGCCATGCCCTGC 871

RESULT 9
 US-09-812-946A-7
 Sequence 7, Application US/08812946A
 Patent No. 6221637

GENERAL INFORMATION:
 APPLICANT: Tsuneaki HIDA et al.
 TITLE OF INVENTION: XANTHENE DERIVATIVES, THEIR PRODUCTION AND
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wenderoth, Lind & Ponack
 STREET: 805 Fifteenth Street, N.W., #700
 CITY: Washington
 STATE: D.C.
 COUNTY: U.S.A.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb

COMPUTER: IBM Compatible
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: Wordperfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/812,946A
 FILING DATE: March 4, 1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Warren M. Cheek, Jr.
 REGISTRATION NUMBER: 33,367
 REFERENCE/DOCKET NUMBER:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-8850
 TELEFAX:
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 29 bases
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Other nucleic acid, Synthetic DNA
 US-08-812-946A-7

Query Match 60.0%; Score 18; DB 3; Length 29;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 13 TCGGAGGGATAGGTGC 30
 Db 8 TCAGGAGGGATAGGTGC 25

RESULT 10
 US-09-270-767-1027
 Sequence 1027, Application US/09270767

GENERAL INFORMATION:
 APPLICANT: Homburger et al.
 TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 FILE REFERENCE: File Reference: 7326-094
 CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17
 NUMBER OF SEQ ID NOS: 625117
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1027
 LENGTH: 268
 TYPE: DNA
 ORGANISM: *Drosophila melanogaster*
 US-09-270-767-1027

Query Match 60.0%; Score 18; DB 4; Length 268;
 Best Local Similarity 80.8%; Pred. No. 26;
 Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 5 GGATCTGTCAGGAGCATAGGTGC 30
 Db 114 GGATCTGTCAGGAGCATAGGTGC 139

RESULT 11
 US-09-270-767-16309
 Sequence 16309, Application US/09270767
 Patent No. 6703491

GENERAL INFORMATION:
 APPLICANT: Homburger et al.
 TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 FILE REFERENCE: File Reference: 7326-094
 CURRENT APPLICATION NUMBER: US/09/270,767
 CURRENT FILING DATE: 1999-03-17

```

; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 16309
; LENGTH: 268
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-167-16309

Query Match Score 18; DB 4; Length 268;
Best Local Similarity 80.8%; Pred. No. 26;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Software: PatentIn Ver. 2.0
SEQ ID NO: 15260
LENGTH: 1498

RESULT 12
US-09-270-167-2215
; Sequence 2215, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 73261094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 15260
LENGTH: 1498

Query Match Score 18%; DB 4; Length 1498;
Best Local Similarity 80.8%; Pred. No. 40;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Organism: Drosophila melanogaster
US-09-270-767-15260

RESULT 13
US-09-270-167-2215
; Sequence 2215, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 73261094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 15260
LENGTH: 1498

Query Match Score 18%; DB 4; Length 746;
Best Local Similarity 80.8%; Pred. No. 34;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Organism: Drosophila melanogaster
US-09-270-167-2215

Query Match Score 18%; DB 4; Length 746;
Best Local Similarity 80.8%; Pred. No. 34;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Organism: Drosophila melanogaster
US-09-270-167-2215

RESULT 14
US-09-270-767-17497
; Sequence 17497, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 73261094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 17497
LENGTH: 746
TYPE: DNA
Organism: Drosophila melanogaster
US-09-270-767-17497

Query Match Score 18%; DB 4; Length 746;
Best Local Similarity 80.8%; Pred. No. 34;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Organism: Drosophila melanogaster
US-09-270-767-15260/c

RESULT 15
US-09-181-339-4
; Sequence 4, Application US/09181339
; Patent No. 6610827
; GENERAL INFORMATION:
; APPLICANT: Forsayeth, John R.
; APPLICANT: Zhao, Byron
; APPLICANT: Chavez, Raymond C.
; TITLE OF INVENTION: POTASSIUM CHANNEL SUBUNIT POLYPEPTIDE
; FILE REFERENCE: 5865-0033-30
; CURRENT APPLICATION NUMBER: US/09/181,339
; CURRENT FILING DATE: 1998-10-28
; PRIORITY NUMBER: US 60/063,450
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PeptSeq for Windows Version 3.0
; SEQ ID NO: 4
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: CDS
; NAME/KEY: CDS
; LOCATION: (2)...(772)
US-09-181-339-4

Query Match Score 17.8%; DB 4; Length 774;
Best Local Similarity 75.9%; Pred. No. 43;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Organism: Homo sapiens
US-09-181-339-4

RESULT 16
US-09-221-017B-726/C
; Sequence 726, Application US/09221017B
; Patent No. 644799
; GENERAL INFORMATION:
; APPLICANT: Roots, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEES: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; ZIP: 94301
; COUNTRY: US

Query Match Score 18%; DB 4; Length 774;
Best Local Similarity 75.9%; Pred. No. 43;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Organism: P. gingivalis
US-09-221-017B-726/C

Query Match Score 18%; DB 4; Length 30;
Best Local Similarity 75.9%; Pred. No. 43;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Organism: P. gingivalis
US-09-221-017B-726/C

Query Match Score 18%; DB 4; Length 30;
Best Local Similarity 75.9%; Pred. No. 43;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Organism: P. gingivalis
US-09-221-017B-726/C

Query Match Score 18%; DB 4; Length 30;
Best Local Similarity 75.9%; Pred. No. 43;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Organism: P. gingivalis
US-09-221-017B-726/C

```

```

COUNTRY: USA ; LENGTH: 1758 ;
ZIP: 94304-1018 ; TYPE: DNA ;
COMPUTER READABLE FORM: ; ORGANISM: Pseudomonas aeruginosa
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Moncay, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-200021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 726:
SEQUENCE CHARACTERISTICS:
LENGTH: 1974 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORYPHYROMONAS GINGIVALIS
NAME/KEY: misc_feature
LOCATION: 1..1974
US-09-221-01-B-726

Query Match 59.3%; Score 17.8; DB 4; Length 1974;
Best Local Similarity 75.9%; Pred. No. 54; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TCGGGGATTTGTAGGGACGGATAAGGCTG 29
Db 983 TCGGGGAGCTTGTGGCAAGGAATACGCCG 955

RESULT 17
US-09-252-991A-7065/C ; LENGTH: 136 ;
Sequence 7065, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEARIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196-136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIORITY APPLICATION NUMBER: US/07/4,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7065

Query Match 58.0%; Score 17.4; DB 4; Length 3186;
Best Local Similarity 77.8%; Pred. No. 94; Mismatches 6; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GAGGATCTTGTAGGGAGGATAGGCTG 29
Db 125 GAGGATCCCGTCATGACCGATAACCTG 99

RESULT 18
US-09-252-991A-7009 ; LENGTH: 136 ;
Sequence 7009, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196-136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIORITY APPLICATION NUMBER: US/07/4,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7009

Query Match 58.0%; Score 17.4; DB 4; Length 2910;
Best Local Similarity 77.8%; Pred. No. 92; Mismatches 6; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GAGGATCTTGTAGGGAGGATAGGCTG 29
Db 19 GAGGATCCCGTCATGACCGATAACCTG 45

RESULT 19
US-09-252-991A-7120/C ; LENGTH: 136 ;
Sequence 7120, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196-136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIORITY APPLICATION NUMBER: US/07/4,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7120
LENGTH: 3186
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7120

Query Match 58.0%; Score 17.4; DB 4; Length 3186;
Best Local Similarity 77.8%; Pred. No. 94; Mismatches 6; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GAGGATCTTGTAGGGAGGATAGGCTG 29
Db 19 GAGGATCCCGTCATGACCGATAACCTG 45

```

Db 3063 GAGGATCCCGTCACTGGATAACCTG 3 037

RESULT 20
 US-09-513-999C-25301
 ; Sequence 25301, Application US/09513999C
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; DUCIERT, A.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
 ; FILE REFERENCE: 59 US2 REG
 ; CURRENT APPLICATION NUMBER: US/09/513,999C
 ; CURRENT FILING DATE: 2000-02-24
 ; PRIORITY APPLICATION NUMBER: US 60/122,487
 ; PRIOR FILING DATE: 1999-02-26
 ; NUMBER SEQ ID NOS: 36681
 ; SOFTWARE: Patent .pm
 ; SEQ ID NO 25301
 ; LENGTH: 351
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 148
 ; OTHER INFORMATION: s=g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 345
 ; OTHER INFORMATION: n=a, g, c or t
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 346
 ; OTHER INFORMATION: s=g or c

US-09-513-999C-25301
 Query Match 57.3%; Score 17.2; DB 4; Length 351;
 Best Local Similarity 73.3%; Pred. No. 68; Indels 0; Gaps 0;
 Matches 22; Conservative 0; Mismatches 8;
 Qy 1 TCGAGGATTGTCAAGGGATAGGCTGC 30
 Db 10 TCGAAGATTGTTCAAGGGTGGGGCG 39

RESULT 21
 US-09-270-767-4139
 ; Sequence 4139, Application US/09270767
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER SEQ ID NOS: 65117
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4139
 ; LENGTH: 534
 ; TYPE: DNA
 ; ORGANISM: Drosophila melanogaster
 ; OTHER INFORMATION: n means any nucleotide

US-09-270-767-4139
 Query Match 57.3%; Score 17.2; DB 4; Length 534;
 Best Local Similarity 86.4%; Pred. No. 75; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 3;

Qy 3 GAGGATCTGTCAGGAGGATA 24
 ; OTHER INFORMATION: n means any nucleotide
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER SEQ ID NOS: 65117
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4139
 ; LENGTH: 534
 ; TYPE: DNA
 ; ORGANISM: Drosophila melanogaster
 ; OTHER INFORMATION: n means any nucleotide

Db 204 GAAGTCTCTGTTAGGGAGCGATA 225

RESULT 22
 US-09-270-767-19421
 ; Sequence 19421, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 65117
 ; SEQ ID NO 19421
 ; LENGTH: 534
 ; SOFTWARE: PatentIn Ver. 2.0
 ; TYPE: DNA
 ; ORGANISM: Drosophila melanogaster
 ; FEATURE:
 ; OTHER INFORMATION: n means any nucleotide
 ; SEQ ID NO 19421

US-09-270-767-19421
 Query Match 57.3%; Score 17.2; DB 4; Length 534;
 Best Local Similarity 86.4%; Pred. No. 75; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 3;

Qy 3 GAGGATCTGTCAGGAGGATA 24
 Db 204 GAAGTCTCTGTTAGGGAGCGATA 225

RESULT 23
 US-09-489-039A-3420
 ; Sequence 3420, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 3420
 ; LENGTH: 939
 ; TYPE: DNA
 ; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-3420
 Query Match 57.3%; Score 17.2; DB 4; Length 939;
 Best Local Similarity 73.3%; Pred. No. 87; Indels 0; Gaps 0;
 Matches 22; Conservative 0; Mismatches 8;

Qy 1 TCGAGGATCTGTCAGGAGGATAAGCTGC 30

Db 566 TCTATGACCTGTCAGGAGTGGCCGC 595

RESULT 24
 US-09-741-238-11/C
 ; Sequence 11, Application US/09741238
 ; Patent No. 6706867
 ; GENERAL INFORMATION:
 ; APPLICANT: Lorenz, Matthias
 ; TITLE OF INVENTION: DNA Array Sequence Selection
 ; FILE REFERENCE: NIH-05076
 ; CURRENT APPLICATION NUMBER: US/09/741,238
 ; CURRENT FILING DATE: 2001-08-20
 ; NUMBER OF SEQ ID NOS: 29

Qy 3 GAGGATCTGTCAGGAGGATA 24

SOFTWARE: PatentIn version 3.0
 SEQ ID NO 11
 LENGTH: 1492
 TYPE: DNA
 ORGANISM: Mus musculus
 US-09-741-238-11

Query Match Score 17.2; DB 4; Length 1492;
 Best Local Similarity 73.3%; Pred. No. 1e+02;
 Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

RESULT 25
 US-09-600-991-3/C
 Sequence 3, Application US/09600991
 GENERAL INFORMATION:
 APPLICANT: MEDICO, Enzo
 APPLICANT: MICHELI, Paolo
 APPLICANT: COLESI, Chiara
 APPLICANT: CASELLI, Gianfranco
 APPLICANT: COMOGLIO, Paolo
 TITLE OF INVENTION: RECOMBINANT PROTEINS DERIVED FROM HGF AND MSP
 FILE REFERENCE: 0471-0162P
 CURRENT APPLICATION NUMBER: US/09/600,991
 CURRENT FILING DATE: 2001-08-20
 NUMBER OF SEQ ID NOS: 22
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 3
 LENGTH: 1692
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Metron F-1 DNA coding sequence

RESULT 26
 US-09-600-991-21/C
 Sequence 21, Application US/09600991
 GENERAL INFORMATION:
 APPLICANT: MEDICO, Enzo
 APPLICANT: MICHELI, Paolo
 APPLICANT: COLESI, Chiara
 APPLICANT: CASELLI, Gianfranco
 APPLICANT: COMOGLIO, Paolo
 TITLE OF INVENTION: RECOMBINANT PROTEINS DERIVED FROM HGF AND MSP
 FILE REFERENCE: 0471-0162P
 CURRENT APPLICATION NUMBER: US/09/600,991
 CURRENT FILING DATE: 2001-08-20
 NUMBER OF SEQ ID NOS: 22
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 21
 LENGTH: 1709
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Metron F-1

Query Match Score 17.2; DB 1; Length 2021;
 Best Local Similarity 73.3%; Pred. No. 1e+02;
 Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 TCGAGGATCTGTCAAGGCGATAAGCTGC 30
 Db 705 TCAGGAGACTTGGCCGCTCGAAGGGTGC 676

RESULT 27
 US-07-882-925A-3/C
 Sequence 3, Application US/07882925A
 GENERAL INFORMATION:
 APPLICANT: Degen, Sandra J. F.
 TITLE OF INVENTION: Gene for a growth factor and its cDNA and
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wood, Herron & Evans, 2700 Carew Tower
 STREET: Wood, Herron & Evans, 2700 Carew Tower
 CITY: Cincinnati
 STATE: Ohio
 COUNTRY: USA
 ZIP: 45202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Macintosh 6.0.3
 SOFTWARE: Microsoft Word 4.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/882,925A
 FILING DATE: 1992-05-14
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Lunn, Gregory
 REGISTRATION NUMBER: 29,945
 REFERENCE/DOCKET NUMBER: CMC 57
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (513) 241-2224
 TELEFAX: (513) 421-7259
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2021 base pairs
 TYPE: NUCLEAR ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 ANTI-SENSE: no
 ORIGINAL SOURCE:
 ORGANISM: human 3p21/p31F15S2
 DEVELOPMENTAL STAGE: fetal
 TISSUE TYPE: liver
 IMMEDIATE SOURCE:
 LIBRARY: cDNA
 CLONE: #19
 POSITION IN GENOME:
 CHROMOSOME SEGMENT: Human 3p21/p31F15S2
 FEATURE:
 IDENTIFICATION METHOD: experimental
 OTHER INFORMATION: This sequence is a variant where two regions were found to be deleted when compared to SEQ ID NO:1.
 PUBLICATION INFORMATION:
 /RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 2021
 US-07-882-925A-3

Query Match Score 17.2; DB 1; Length 2021;
 Best Local Similarity 73.3%; Pred. No. 1e+02;
 Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 TCGAGGATCTGTCAAGGCGATAAGCTGC 30
 Db 660 TCAGGAGACTTGGCCGCTCGAAGGGTGC 631

RESULT 28
US-08-184-012C-3/c
Sequence 3, Application US/08184012C
Patent No. 5,606029
GENERAL INFORMATION:
APPLICANT: Degen, Sandra J. F.
TITLE OF INVENTION: Gene for a growth factor and its cDNA and recombinant proteins derived from HGF and MSP
FILE REFERENCE: 0471-0162P
CURRENT APPLICATION NUMBER: US/09/600,991
CURRENT FILING DATE: 2001-08-20
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.1
SEQ ID NO 19
LENGTH: 2136
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2136)
OTHER INFORMATION:
US-09-600-991-19

Query Match 57.3%; Score 17.2; DB 4; Length 2136;
Best Local Similarity 73.3%; Pred. No. 1..1e+02;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 TCGAGGATCTTGTCAAGGGATAGGCTGC 30
Db 695 TCGAGGAACTTGCCGGCTCGAAGGGTGC 666

RESULT 30
US-09-601-040A-11/C
Sequence 11, Application US/09601040A
Patent No. 6,730657
GENERAL INFORMATION:
APPLICANT: MEDICO, Enzo et al.
TITLE OF INVENTION: Recombinant Proteins from HGF and MSP
FILE REFERENCE: 0471-0161P
CURRENT APPLICATION NUMBER: US/09/601,040A
CURRENT FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn version 3.0
SEQ ID NO 11
LENGTH: 2136
TYPE: DNA
ORGANISM: Homo sapiens
US-09-601-040A-11

Query Match 57.3%; Score 17.2; DB 4; Length 2136;
Best Local Similarity 73.3%; Pred. No. 1..1e+02;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 TCGAGGATCTTGTCAAGGGATAGGCTGC 30
Db 695 TCGAGGAACTTGCCGGCTCGAAGGGTGC 666

RESULT 31
US-07-882-925A-4/c
Sequence 4, Application US/07882925A
Patent No. 5315000
GENERAL INFORMATION:
APPLICANT: Degen, Sandra J. F.
TITLE OF INVENTION: Gene for a growth factor and its cDNA and recombinant proteins derived from HGF and MSP
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gregory Lunn
STREET: Wood, Herron & Evans, 2700 Carew Tower
CITY: Cincinnati
STATE: Ohio
COUNTRY: USA
ZIP: 45202

RESULT 29
US-09-600-991-19/c
Sequence 19, Application US/09600991
Patent No. 6551991
GENERAL INFORMATION:

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Macintosh 6.0.3
 SOFTWARE: Microsoft Word 4.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/882,925A
 FILING DATE: 19920514
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Lunn, Gregory
 REGISTRATION NUMBER: 29,945
 REFERENCE/DOCKET NUMBER: CMC 57
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (513) 241-2324
 TELEFAX: (513) 421-7269
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2188 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 ANTI-SENSE: no
 ORIGINAL SOURCE:
 ORGANISM: mouse
 STRAIN: C57BL/6
 DEVELOPMENTAL STAGE: adult
 TISSUE TYPE: liver
 IMMEDIATE SOURCE:
 LIBRARY: cDNA
 CLONE: ML5-2
 POSITION IN GENOME:
 CHROMOSOME SEGMENT: mouse 9, Hgf1 locus
 MAP POSITION: Trf-Gnai-2-Hgf1-Cck
 FEATURE:
 IDENTIFICATION METHOD: experimental
 PUBLICATION INFORMATION:
 RELEVANT RESIDUES IN SEQ ID NO: 4: 1 TO 2188
 US-07-882,925A-4

Query Match Score 57.3%; Pred. No. 1.1e+02; Length 2188;
 Best Local Similarity 73.3%; Pred. No. 1.1e+02;
 Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

RESULT 32
 US-09-184-012C-4/C
 Sequence 4, Application US/08184012C
 Patent No. 5606029
 GENERAL INFORMATION:
 APPLICANT: Degen, Sandra J. F.
 TITLE OF INVENTION: Gene for a growth factor and its cDNA and
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Gregory Lunn
 STREET: Wood, Herron & Evans, 2700 Carew Tower
 CITY: Cincinnati
 STATE: Ohio
 COUNTY: USA
 ZIP: 45202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Macintosh 7.5.2
 SOFTWARE: Microsoft Word 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/666,082B
 FILING DATE: 19-JUN-1996
 CLASSIFICATION: 536
 PRIORITY APPLICATION DATA:

FILING DATE: 1/18/94
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Lunn, Gregory
 REGISTRATION NUMBER: 29,945
 REFERENCE/DOCKET NUMBER: CMC 57
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (513) 241-2324
 TELEFAX: (513) 421-7269
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2188 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 ANTI-SENSE: no
 ORIGINAL SOURCE:
 ORGANISM: mouse
 STRAIN: C57BL/6
 DEVELOPMENTAL STAGE: adult
 TISSUE TYPE: liver
 IMMEDIATE SOURCE:
 LIBRARY: cDNA
 CLONE: ML5-2
 POSITION IN GENOME:
 CHROMOSOME SEGMENT: mouse 9, Hgf1 locus
 MAP POSITION: Trf-Gnai-2-Hgf1-Cck
 FEATURE:
 IDENTIFICATION METHOD: experimental
 PUBLICATION INFORMATION:
 RELEVANT RESIDUES IN SEQ ID NO: 4: 1 TO 2188
 US-08-184-012C-4

Query Match Score 57.3%; Pred. No. 1.1e+02; Length 2188;
 Best Local Similarity 73.3%; Pred. No. 1.1e+02;
 Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 TCGAGGATCTTGACGGATGGCTGC 30
 Db 651 TCTAGGAACTTTCAGGCTGGAAAGGCTGC 622

RESULT 33
 US-08-666-082B-2/C
 Sequence 2, Application US/08666082B
 Patent No. 5916770
 GENERAL INFORMATION:
 APPLICANT: YOSHIKAWA, WATARU
 APPLICANT: SHIMONISHI, MANABU
 APPLICANT: IWAMOTO, JUNKO
 APPLICANT: TAKEHARA, TOYIRO
 APPLICANT: HAGIYA, MICHIO
 TITLE OF INVENTION: MACROPHAGE STIMULATING PROTEIN VARIANT AND
 TITLE OF INVENTION: METHOD FOR PRODUCING THE SAME
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LEDIG, VOIT & MAYER, LTD.
 STREET: TWO PRUDENTIAL PLAZA, SUITE 4900
 CITY: CHICAGO
 STATE: IL
 COUNTRY: US
 ZIP: 60601-6780
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/666,082B
 FILING DATE: 19-JUN-1996
 CLASSIFICATION: 536
 PRIORITY APPLICATION DATA:

```

APPLICATION NUMBER: JP 7-153309
FILING DATE: 20-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ROBERT F. GREEN
REGISTRATION NUMBER: 27555
REFERENCE/DOCKET NUMBER: 73843
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2216 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Human
FEATURE:
NAME/KEY: sig peptide
LOCATION: 1..54
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2133
US-08-666-082B-2

Query Match      57.3% Score
Best Local Similarity 73.3% Pred.
Matches 22; Conservative 0; Mis
Qy   1 TCGAGGATCTGTCAAGGGATAGC
Db   695 TCAGGAACTTGCCCGTCGAAGG

```

RESULT 34

US-07-882-925A-1/c

Sequence 1, Application US/07882925A
Patent No. 5215000.

GENERAL INFORMATION:
APPLICANT: Degen, Sandra J. F.
TITLE OF INVENTION: Gene for a growth protein
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADRESSEE: Gregory Lunn
STREET: Wood, Herron & Evans, 27
CITY: Cincinnati
STATE: Ohio
COUNTRY: USA
ZTF: 45202

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/882,92
FILING DATE: 1992/05/14
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Gregory
REGISTRATION NUMBER: 29,945
REFERENCE/DOCKET NUMBER: CMC 57
TELECOMMUNICATION INFORMATION:
TELEPHONE: (513) 241-2324
TELEFAX: (513) 421-7269
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2219 base pairs
TYPE: NUCLEAR ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ANTI SENSE: no
ORIGINAL SOURCE:

```

ORGANISM: human
DEVELOPMENTAL STAGE: fetal
TISSUE TYPE: liver
IMMEDIATE SOURCE:
LIBRARY: CDNA
CLONE: #33
POSITION IN GENOME: POSITION IN GENOME: human 3p21/D3F15S2
CHROMOSOME/SEGMENT: FEATURE:
FEATURE: IDENTIFICATION METHOD: experimental
OTHER INFORMATION: Includes five polymers at the same level; one of which is mismatched.
OTHER INFORMATION: Other information ID: 619.
OTHER INFORMATION: Polymorphic amino acid.
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 1: FRR
US-07-382-925A-1

Query Match Score 17.2;
Best Local Similarity 73.3%; Pred. No. 1.
Matches 22; Conservative 0; Mismatches 619.
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 1: FRR
US-07-382-925A-1

      RESULT 35
      US-07-382-925A-2/c
      Sequence 2, Application US/07882925A
      Patent No. 5313000
      GENERAL INFORMATION:
      APPLICANT: Degen, Sandra J. F.
      TITLE OF INVENTION: Gene for a growth factor
      TITLE OF INVENTION: protein
      NUMBER OF SEQUENCES: 7
      CORRESPONDENCE ADDRESS:
      ADDRESSEE: Gregory Lunn
      STREET: Wood, Herron & Evans, 2700 Carr
      CITY: Cincinnati
      STATE: Ohio
      COUNTRY: USA
      ZIP: 45212
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Diskette, 3.50 inch, 800
      COMPUTER: Apple Macintosh
      OPERATING SYSTEM: Macintosh 6.0.3
      SOFTWARE: Microsoft Word 4.0
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/07/882,925A
      FILING DATE: 19920314
      CLASSIFICATION: 530
      ATTORNEY/AGENT INFORMATION:
      NAME: Lunn, Gregory
      REGISTRATION NUMBER: 29,945
      REFERENCE/DOCKET NUMBER: CMC 57
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (513) 241-2324
      TELEFAX: (513) 421-7269
      INFORMATION FOR SEQ ID NO: 2:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 2219 base pairs
      TYPE: NUCLEAR ACID
      STRANDEDNESS: single
      TOPOLOGY: linear
      MOLECULE TYPE: cDNA to mRNA
      ANTI -SENSE: no
      ORIGINAL SOURCE:
      ORGANISM: human
      DEVELOPMENTAL STAGE: fetal
      TISSUE TYPE: liver
      LIBRARY: CDNA

```

CLONE: #33
 POSITION IN GENOME: human 3p21/D3F15S2
 CHROMOSOME/SEGMENT: human 3p21/D3F15S2

FEATURE: IDENTIFICATION METHOD: experimental
 OTHER INFORMATION: Includes five polymorphisms at the nucleotide level; one of which results in an amino acid substitution (nucl. position 619). Sequence ID NO:1: contains the identical sequence with other information: polymorphic amino acid.

OTHER INFORMATION: RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 2219
 US-07-882-925A-2

RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2219
 US-07-882-925A-2

Query Match 57.3%; Score 17.2; DB 1; Length 2219;
 Best Local Similarity 73.3%; Pred. No. 1.1e+02;
 Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 TCGGGGATCTGTCAAGGACGATAGGCTGC 30
 Db 679 TCGGGAAACTTGCCTGGCTGAAGGGTC 650

RESULT 37
 US-08-184-012C-2/c
 Sequence 2, Application US/08184012C
 Patent No. 5606029

GENERAL INFORMATION:
 APPLICANT: Degen, Sandra J. F.
 TITLE OF INVENTION: Gene for a growth factor and its cDNA and protein
 NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Gregory Lunn
 STREET: Wood, Herron & Evans, 2700 Carew Tower
 CITY: Cincinnati
 STATE: Ohio
 ZIP: 45202

COMPUTER READABLE FORM:
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Macintosh 7.5.2
 SOFTWARE: Microsoft Word 5.1

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/184,012C
 FILING DATE: 1/18/94
 CLASSIFICATION: 516

ATTORNEY/AGENT INFORMATION:
 NAME: Lunn, Gregory
 REGISTRATION NUMBER: 29,945
 REFERENCE/DOCKET NUMBER: CMC 57

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (513) 421-2324
 TELEFAX: (513) 421-7269

REGISTRATION NUMBER: 29,945
 REFERENCE/DOCKET NUMBER: CMC 57

SEQUENCE CHARACTERISTICS:
 LENGTH: 2219 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 ANTI-SENSE: no
 ORIGINAL SOURCE:
 ORGANISM: human
 DEVELOPMENTAL STAGE: fetal
 TISSUE TYPE: liver
 IMMEDIATE SOURCE:
 LIBRARY: cDNA
 CLONE: #33

POSITION IN GENOME: human 3p21/D3F15S2
 CHROMOSOME/SEGMENT: human 3p21/D3F15S2

FEATURE: IDENTIFICATION METHOD: experimental
 OTHER INFORMATION: Includes five polymorphisms at the nucleotide level; one of which results in an amino acid substitution (nucl. position 619). Sequence ID NO:1: contains the identical sequence with other information: polymorphic amino acid.

Patent No. 5606029

OTHER INFORMATION: polymorphic amino acid.
 PUBLICATION INFORMATION:
 RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2219
 US-08-184-012C-2

Query Match 57.3%; Score 17.2; DB 1;
 Best Local Similarity 73.3%; Prod. No. 1.1e+02;
 Matches 22; Conservative 0; Mismatches 8;
 Indels 0; Gaps 0;
 Qy 1 TCGAGGATTCGGTCAAGGAGATAAGGCTG 30
 Db 679 TCGAGGAACCTGCCGGCTGAAGGGTGC 650

RESULT 38

US-08-177-1/c

Sequence 1, Application US/06334177

Patent No. 5636086

GENERAL INFORMATION:

APPLICANT: Avraham, Hava Karsenty

TITLE OF INVENTION: Methods and Kits Using Macrophage Stimulating Protein

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESS SEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin. (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/334,177

FILING DATE:

CLASSIFICATION: 435

PRIORITY APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Marschang, Diane L.

REGISTRATION NUMBER: 35,600

REFERENCE/DOCKET NUMBER: 912

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-5416

TELEFAX: 415/952-9881

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2232 bases

TYPE: nucleic acid

STRANDEDNESS: Single

TOPOLOGY: linear

US-08-177-1

Query Match 57.3%; Score 17.2; DB 1;
 Best Local Similarity 73.3%; Prod. No. 1.1e+02;
 Matches 22; Conservative 0; Mismatches 8;
 Indels 0; Gaps 0;

Qy 1 TCGAGGATTCGGTCAAGGAGATAAGGCTG 30
 Db 705 TCGAGGAACCTGCCGGCTGAAGGGTGC 676

RESULT 39

PC-US95-13830-1/c

Sequence 1, Application PC/TUSS513830

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

APPLICANT: New England Deaconess Hospital Corp.

TITLE OF INVENTION: Methods and Kits Using Macrophage Stimulating Protein
 TITLE OF INVENTION: Protein
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPat-in (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/13830
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Marschang, Diane L.
 REGISTRATION NUMBER: 35,600
 REFERENCE/DOCKET NUMBER: P0912PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-5416
 TELEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2232 base pairs
 TYPE: Nucleic Acid
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 PCT-US95-13830-1

Query Match 57.3%; Score 17.2; DB 5;
 Best Local Similarity 73.3%; Prod. No. 1.1e+02;
 Matches 22; Conservative 0; Mismatches 8;
 Indels 0; Gaps 0;

Qy 1 TCGAGGATTCGGTCAAGGAGATAAGGCTG 30
 Db 705 TCGAGGAACCTGCCGGCTGAAGGGTGC 676

RESULT 40

US-07-882-925A-7/c

Sequence 7, Application US/07882925A
 Patent No. 5315000

GENERAL INFORMATION:
 APPLICANT: Degen, Sandra J. F.
 TITLE OF INVENTION: Gene for a growth factor and its cDNA and
 NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Gregory Lunn
 STREET: Wood, Herron & Evans, 2700 Carew Tower
 CITY: Cincinnati
 STATE: Ohio
 COUNTRY: USA
 ZIP: 45202

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Macintosh 6.0.3
 SOFTWARE: Microsoft Word 4.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/882,925A
 FILING DATE: 19920514
 CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
 NAME: Lunn, Gregory
 REGISTRATION NUMBER: 29,945
 REFERENCE/DOCKET NUMBER: CMC 57

TELECOMMUNICATION INFORMATION:
TELEPHONE: (513) 241-2324
TELEFAX: (513) 421-7269
SEQUENCE CHARACTERISTICS:
LENGTH: 2262 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: Identical to sequence ID NO: 1; with 5' and 3'
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: human
DEVELOPMENTAL STAGE: fetal
TISSUE TYPE: liver
IMMEDIATE SOURCE:
LIBRARY: cDNA
CLONE: #33 including 5' and 3' adaptors
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 7: FROM 1 TO 2262
US-07-882-92A-7

Query Match 57.3%; Score 172; DB 1; Length 2262;
Best Local Similarity 73.3%; Pred. No 1.1e+02;
Matches 22; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 1 TCGAGGATTTGTAGGGACGATAGGCTGC 30
Db 710 TCGAGGAACCTTGCCCGGCTGAAAGGGTGC 681

Search completed: January 7, 2005, 11:00:37
Job time : 56.2105 Secs

Copyright (c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6

OM nucleic - nucleic search, using SW model

Run on: January 7, 2005, 11:00:44 ; Search time 1275.26 Seconds
 134.979 Million cell updates/sec

Title: US-09-786-502A-8

Perfect score: 30

Sequence: tcgagatgtcgaggagtcatacgctgc 30

Scoring table: IDENTITY_NUC

Gapext 1.0

Searched: 4293498 seqs, 2868903791 residues

Total number of hits satisfying chosen parameters: 8586996

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 10%

Listing first 45 summaries

Database : Published Applications NA:
 1: /cgn2_6/_ptodata/2/pubnna/us07_PUBCOMB.seq,*
 2: /cgn2_6/_ptodata/2/pubnna/pct_NEW_PUB.seq,*
 3: /cgn2_6/_ptodata/2/pubnna/us06_NEWPUB.seq,*
 4: /cgn2_6/_ptodata/2/pubnna/us07_NEWPUB.seq,*
 5: /cgn2_6/_ptodata/2/pubnna/fctus_PUBCOMB.seq,*
 6: /cgn2_6/_ptodata/2/pubnna/fctus_PUBCOMB.seq,*
 7: /cgn2_6/_ptodata/2/pubnna/us07_NEWPUB.seq,*
 8: /cgn2_6/_ptodata/2/pubnna/us08_NEWPUB.seq,*
 9: /cgn2_6/_ptodata/2/pubnna/us09_NEWPUB.seq,*
 10: /cgn2_6/_ptodata/2/pubnna/us09B_PUBCOMB.seq,*
 11: /cgn2_6/_ptodata/2/pubnna/us09C_PUBCOMB.seq,*
 12: /cgn2_6/_ptodata/2/pubnna/us09_NEWPUB.seq,*
 13: /cgn2_6/_ptodata/2/pubnna/us10A_PUBCOMB.seq,*
 14: /cgn2_6/_ptodata/2/pubnna/us10B_PUBCOMB.seq,*
 15: /cgn2_6/_ptodata/2/pubnna/us10C_PUBCOMB.seq,*
 16: /cgn2_6/_ptodata/2/pubnna/us10D_PUBCOMB.seq,*
 17: /cgn2_6/_ptodata/2/pubnna/us10E_PUBCOMB.seq,*
 18: /cgn2_6/_ptodata/2/pubnna/us10_NEWPUB.seq,*
 19: /cgn2_6/_ptodata/2/pubnna/us11_NEWPUB.seq,*
 20: /cgn2_6/_ptodata/2/pubnna/us60_NEWPUB.seq,*
 21: /cgn2_6/_ptodata/2/pubnna/us60_NEWPUB.seq,*

RESULT 1
 US-08-940-544-2

; Sequence 2, Application US/08940544B
 ; Publication No. US20020018783A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SADELLAIN, MICHEL
 ; APPLICANT: CHEUNG, NAI-KONG V.
 ; APPLICANT: KRAUSE, ANJA
 ; APPLICANT: GUO, HONG-FEN
 ; TITLE OF INVENTION: FUSION PROTEINS OF A SINGLE CHAIN ANTIBODY AND CD28 AND
 ; TITLE OF INVENTION: FUSION PROTEINS OF A SINGLE CHAIN ANTIBODY AND CD28 AND
 ; FILE REFERENCE: MSK_P-035-US
 ; CURRENT APPLICATION NUMBER: US/08/940,544B
 ; CURRENT FILING DATE: 1997-03-30
 ; EARLIER APPLICATION NUMBER: PCT/US97/04427
 ; EARLIER FILING DATE: 1997-03-20
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 30
 ; TYPE: DNA
 ; ORGANISM: HUMAN
 ; FEATURE:
 ; OTHER INFORMATION: Downstream primer for PCR amplification.
 ; US-08-940-544-2

Query Match Score 100.0%; Score 30; DB 8; Length 30;
 Best Local Similarity 100.0%; Pred. No. 0.0007%;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCGAGATCTTGTCAAGGCGATAAGCTGC 30
 Db 1 TCGAGATCTTGTCAAGGCGATAAGCTGC 30

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	8	US-08-940-544-2
c 2	20.4	68.0	1514	9	US-09-738-546-1
c 3	20.4	68.0	1514	11	US-09-836-544-7
c 4	20.4	68.0	1514	15	US-10-207-655-98
c 5	20.4	68.0	1514	16	US-10-641-643-1316
c 6	20.4	68.0	3803	15	US-10-976-932-1
c 7	20.4	68.0	3804	11	US-09-987-722-257
c 8	20.4	68.0	3804	15	US-10-143-238-1
c 9	20.4	68.0	3805	10	US-09-835-297-3
c 10	20.4	68.0	51365	11	US-09-397-256-8
c 11	20.2	67.3	36	16	US-10-448-256-8
c 12	20	66.7	32	15	US-10-273-762-2

RESULT 2
 US 09-738-546-1/c
 Sequence 1, Application US/09738546
 Patent No. US20020006403A1
 GENERAL INFORMATION:
 APPLICANT: YU, XUE ZHONG
 APPLICANT: ANASSETTI, CLAUDIO
 TITLE OF INVENTION: CD28-SPECIFIC ANTIBODY COMPOSITIONS FOR USE IN METHODS OF IMMUNOSUPPRESSION
 FILE REFERENCE: FHCIC:00705
 CURRENT APPLICATION NUMBER: US/09/738, 546
 CURRENT FILING DATE: 2000-12-14
 PRIORITY NUMBER: 60/170, 857.
 PRIOR FILING DATE: 1999-12-14
 NUMBER OF SEQ ID NOS: 2
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1
 LENGTH: 1514
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-738-546-1

Query Match 68.0%; Score 20.4; DB 9; Length 1514;
 Best Local Similarity 95.5%; Prod. No. 23;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 3
 US-09-836-544-7/c
 Sequence 7, Application US/09836544
 Publication No. US20040072283A1
 GENERAL INFORMATION:
 APPLICANT: The General Hospital Corporation
 TITLE OF INVENTION: Rapid Immunoselection Cloning Method
 FILE REFERENCE: 11-88L
 CURRENT APPLICATION NUMBER: US/09/836, 544
 CURRENT FILING DATE: 2001-04-17
 PRIOR APPLICATION NUMBER: US 07/983, 647
 PRIOR FILING DATE: 1992-12-01
 PRIOR APPLICATION NUMBER: US 07/553, 759
 PRIOR FILING DATE: 1990-07-13
 PRIOR APPLICATION NUMBER: US 07/498, 809
 PRIOR FILING DATE: 1990-03-23
 PRIOR APPLICATION NUMBER: US 07/379, 076
 PRIOR FILING DATE: 1989-07-13
 PRIOR APPLICATION NUMBER: US 07/160, 416
 PRIOR FILING DATE: 1988-02-25
 NUMBER OF SEQ ID NOS: 37
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 7
 LENGTH: 1514
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-836-544-7

Query Match 68.0%; Score 20.4; DB 11; Length 1514;
 Best Local Similarity 95.5%; Prod. No. 23;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 4
 US-10-207-655-98/c
 Sequence 98, Application US/10207655
 Publication No. US2003018592A1
 GENERAL INFORMATION:
 APPLICANT: Ledbetter, Jeffrey A.
 APPLICANT: Hayden-Ledbetter, Martha S.
 TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
 FILE REFERENCE: 390069_401C1
 CURRENT APPLICATION NUMBER: US/10/207, 655
 CURRENT FILING DATE: 2002-07-25
 NUMBER OF SEQ ID NOS: 426
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 98
 LENGTH: 1514
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-207-655-98

Query Match 68.0%; Score 20.4; DB 15; Length 1514;
 Best Local Similarity 95.5%; Prod. No. 23;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 5
 US-10-641-643-1316/c
 Sequence 1316, Application US/10641643
 Publication No. US2004007703A1
 GENERAL INFORMATION:
 APPLICANT: Cocks, Benjamin G.
 APPLICANT: Susan G. Stuart
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
 CORRESPONDENCE ADDRESS: INCYTE PHARMACEUTICALS, INC.
 NUMBER OF SEQUENCES: 1508
 GENE EXPRESSION
 ADDRESS: 3174 PORTER DRIVE
 STREET: 3174 PORTER DRIVE
 CITY: PALO ALTO
 STATE: CALIFORNIA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/641, 643
 FILING DATE: 14-Aug-2003
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: <Unknown>
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Zaller, Karen J.
 REGISTRATION NUMBER: 37, 071
 REFERENCE/DOCKET NUMBER: PA-0001 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 855-0555
 TELEFAX: (650) 845-4166
 INFORMATION FOR SEQ ID NO: 1316:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1514 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: GENBANK
 CLONE: 9338444
 SEQUENCE DESCRIPTION: SEQ ID NO: 1316 :

RESULT 8
US-10-143-238-1/c

Query Match 68.0%; Score 20.4; DB 16; Length 1514;
Best Local Similarity 95.5%; Pred. No. 23; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 1;

Qy 9 CPTGTCAGGAGGGATAGGCTGC 30
Db 766 CGTGTCAAGGAGGGATAGGCTGC 745

RESULT 6
US-10-076-934-1/c

; Sequence 1, Application US/10076934
; Publication No. US20030170232A1
; GENERAL INFORMATION:
; APPLICANT: O'Hara Jr., Richard S.
; TITLE OF INVENTION: DIFFERENTIAL REGULATION OF T CELL SURVIVAL AND PROLIFERATION
; FILE REFERENCE: GNN-030
; CURRENT APPLICATION NUMBER: US/10/143-238
; CURRENT FILING DATE: 2002-05-10
; PRIORITY NUMBER: US 60/290,097
; PRIOR FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ For Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 3804
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (223) . . . (885)

Query Match 68.0%; Score 20.4; DB 15; Length 3804;
Best Local Similarity 95.5%; Pred. No. 23; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 1;

Qy 9 CTTCTCTAGAGGGATAGGCTGC 30
Db 889 CGTGTCAAGGAGGGATAGGCTGC 868

RESULT 9
US-09-835-297-3/c

; Sequence 3, Application US/09835297
; Publication No. US20030086932A1
; GENERAL INFORMATION:
; APPLICANT: Bluestone, Jeffrey
; TITLE OF INVENTION: THAT BIND TO CTLA4 AND USES THEREFOR
; FILE REFERENCE: GNN-014CP
; CURRENT APPLICATION NUMBER: US/09/835,297
; CURRENT FILING DATE: 2001-04-12
; PRIORITY NUMBER: 60/196,851
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 3
; LENGTH: 3806
; TYPE: DNA
; ORGANISM: Homo sapiens

Query Match 68.0%; Score 20.4; DB 10; Length 3806;
Best Local Similarity 95.5%; Pred. No. 23; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 1;

Qy 9 CTTCTCTAGAGGGATAGGCTGC 30
Db 889 CGTGTCAAGGAGGGATAGGCTGC 868

RESULT 10
US-09-997-722-256/c

; Sequence 256, Application US/09997722
; Publication No. US20040072154A1
; GENERAL INFORMATION:
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71171/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/997,722
; CURRENT FILING DATE: 2001-11-30
; PRIORITY NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIORITY NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 257
; LENGTH: 3804
; TYPE: DNA
; ORGANISM: Homo sapiens

Query Match 68.0%; Score 20.4; DB 11; Length 3804;
Best Local Similarity 95.5%; Pred. No. 23; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 1;

Qy 9 CPTGTCAGGAGGGATAGGCTGC 30
Db 889 CGTGTCAAGGAGGGATAGGCTGC 868

```

; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-7171.RMS/DFP
; CURRENT APPLICATION NUMBER: US/09/997,722
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US/09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 256
; LENGTH: 51365
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-997-722-256

Query Match          68.0%;  Score 20.4;  DB 11;  Length 51365;
Best Local Similarity 95.5%;  Pred. No. 26;  Length 51365;
Matches 21;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;
RESULT 13
US-10-437-963-23123/c
Sequence 23123, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; Tissue of Invention: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5322)1.B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO: 23123
; LENGTH: 1329
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_28233C.1
US-10-437-963-23123

Query Match          64.7%;  Score 19.4;  DB 17;  Length 1329;
Best Local Similarity 79.3%;  Pred. No. 64;
Matches 23;  Conservative 0;  Mismatches 6;  Indels 0;  Gaps 0;
RESULT 14
US-10-437-963-23127
Sequence 23127, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; Tissue of Invention: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5322)1.B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14

Query Match          67.3%;  Score 20.2;  DB 16;  Length 36;
Best Local Similarity 88.0%;  Pred. No. 24;  Length 36;
Matches 22;  Conservative 0;  Mismatches 3;  Indels 0;  Gaps 0;
RESULT 12
US-10-273-762-2
Sequence 2, Application US/10273762
; Publication No. US20030147881A1
; GENERAL INFORMATION:
; APPLICANT: GUO, Hong-fen
; APPLICANT: CHEUNG, Mai-kong V
; TITLE OF INVENTION: METHOD FOR PREPARATION OF SINGLE CHAIN ANTIBODIES
; FILE REFERENCE: 676-A-PCT
; CURRENT APPLICATION NUMBER: US/10/273,762
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: 60/7430,396
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 10/097,558
; PRIOR FILING DATE: 2002-03-08

```

```

; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 23127
; LENGTH: 4467
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_28237C.1
US-10-437-963-23127

Query Match: 1 TCGAGGATCTGTCAAGGACGATAAGCTGC 29
Best Local Similarity 79.3%; Pred. No. 68; Mismatches 6; Indels 0; Gaps 0;
Matches 23; Conservative 0; Gaps 0;

RESULT 15
US-10-437-963-28652/c
; Sequence 28652, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 3B-21(5221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 28652
; LENGTH: 4509
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_3322C.1
US-10-437-963-28652

Query Match: 1 TCGAGGATCTGTCAAGGACGATAAGCTGC 29
Best Local Similarity 79.3%; Pred. No. 68; Mismatches 6; Indels 0; Gaps 0;
Matches 23; Conservative 0; Gaps 0;

RESULT 16
US-10-417-375-128
; Sequence 128, Application US/10417375
; Publication No. US20040219528A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001600
; CURRENT APPLICATION NUMBER: US/10/417,375
; CURRENT FILING DATE: 2003-04-15
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: FastSBQ For Windows Version 4.0
; SEQ ID NO 128
; LENGTH: 430442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-417-375-128

Query Match: 1 TGAGGATCTGTCAAGGAGATGGCTG 29
Best Local Similarity 79.3%; Pred. No. 82; Mismatches 6; Indels 0; Gaps 0;
Matches 23; Conservative 0; Gaps 0;

RESULT 17
US-09-986-480-134
; Sequence 134, Application US/0986480
; Publication No. US2003027999A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: PSS00P1
; FILE REFERENCE: PSS00P1
; CURRENT APPLICATION NUMBER: US/09/986,480
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/12788
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/134,068
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 456
; SEQ ID NO 134
; LENGTH: 866
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (20)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (375)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (847)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-986-480-134

Query Match: 1 GAGGATCTGTCAAGGAGATGGCTG 29
Best Local Similarity 81.5%; Pred. No. 96; Mismatches 5; Indels 0; Gaps 0;
Matches 22; Conservative 0; Gaps 0;

RESULT 18
US-10-632-963-46/c
; Sequence 46, Application US/10632983
; Publication No. US20040067518A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2025P1CD1
; CURRENT APPLICATION NUMBER: US/10/632,983
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: 09/716,129
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: PCT/US99/03939
; PRIOR FILING DATE: 1999-02-24
; PRIOR APPLICATION NUMBER: 60/076,053
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,057
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,052
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,054
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,051
; PRIOR APPLICATION NUMBER:

```



```

; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 441
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-260-238-441

Query Match 62.7%; Score 18.8; DB 15;
Best Local Similarity 76.7%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TCGAGGATCTGTCAGGAGGCATAAGGCCATTAGGCTGC 30
Db 224 TCGTCGATCTGTCAGGCCGATATTCTGC 253

RESULT 25
US-10-369-493-38079
; Sequence 38079, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 4734
; SEQ ID NO 38079
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-38079

Query Match 62.7%; Score 18.8; DB 15;
Best Local Similarity 76.7%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TCGAGGATCTGTCAGGAGGCATAAGGCCATTAGGCTGC 30
Db 224 TCGTCGATCTGTCAGGCCGATATTCTGC 253

RESULT 26
US-10-369-493-38204
; Sequence 38204, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 4734
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-38204

Query Match 62.7%; Score 18.8; DB 15;
Best Local Similarity 76.7%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TCGAGGATCTGTCAGGAGGCATAAGGCCATTAGGCTGC 30
Db 224 TCGTCGATCTGTCAGGCCGATATTCTGC 253

RESULT 24
US-10-369-493-35152
; Sequence 35152, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 4734
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-35152

Query Match 62.7%; Score 18.8; DB 15;
Best Local Similarity 76.7%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TCGAGGATCTGTCAGGAGGCATAAGGCCATTAGGCTGC 30
Db 224 TCGTCGATCTGTCAGGCCGATATTCTGC 253

```

RESULT 27
US-10-437-963-48174/c
Sequence 48174, Application US/10437963
Publication No. US20040123343A1.
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With Title of Invention: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 3B-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 48174
LENGTH: 815
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE: OTHER INFORMATION: Clone ID: PAT_MRT4530_50875C.1
US-10-437-963-48174
Query Match 62.7%; Score 18.8; DB 17; Length 815;
Best Local Similarity 90.9%; Pred. No. 1.2e+02; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 2;
Qy 2 CGAGGATCTGTCAAGGAGGAT 23
Db 670 CGAGGATCTGGCAAGGAGGAT 649
RESULT 30
US-10-2/3-762-7/c
Sequence 7, Application US/10273762
Publication No. US20030147831A1
GENERAL INFORMATION:
APPLICANT: CHEUNG, Nai-Kong V
APPLICANT: GUO, Hong-fen
TITLE OF INVENTION: METHOD FOR PREPARATION OF SINGLE CHAIN ANTIBODIES
FILE REFERENCE: 676-A-PCT
CURRENT APPLICATION NUMBER: US/10/273,762
CURRENT FILING DATE: 2002-10-17
PRIORITY APPLICATION NUMBER: 60/330,396
PRIORITY FILING DATE: 2001-10-17
PRIORITY APPLICATION NUMBER: 10/097,558
PRIORITY FILING DATE: 2002-03-08
PRIORITY APPLICATION NUMBER: PCT/US01/32565
PRIORITY FILING DATE: 2001-10-18
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin version 3.1
SEQ ID NO 7
LENGTH: 48
TYPE: DNA
ORGANISM: Homo Sapiens
US-10-2/3-762-7
Query Match 61.3%; Score 18.4; DB 15; Length 48;
Best Local Similarity 95.0%; Pred. No. 1.6e+02; Mismatches 1;
Matches 19; Conservative 0; Gaps 0;
Qy 11 TCGCAGGAGGATAGGCTGC 30
Db 29 TGCCAGGAGGATAGGCTGC 10
RESULT 31
US-10-172-119-212/C
Sequence 212, Application US/10172118
Publication No. US20030224374A1
GENERAL INFORMATION:

RESULT 28
US-10-369-493-44430/c
Sequence 44430, Application US/10369493
Publication No. US20030233675A1.
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 3B-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIORITY APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 44430
LENGTH: 1431
TYPE: DNA
ORGANISM: Rhodopseudomonas palustris
US-10-369-493-44430
Query Match 62.7%; Score 18.8; DB 15; Length 1431;
Best Local Similarity 90.9%; Pred. No. 1.2e+02; Mismatches 2;
Matches 20; Conservative 0; Gaps 0;
Qy 1 TCGGGATCTGTCAAGGAGGA 22
Db 638 TCAGGGATCTGTGGGATCGGA 617
RESULT 29
US-10-437-963-67598

RESULT 32
 US-10-342-887-212/C
 ; Sequence 212, Application US/10342887
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yudong
 ; APPLICANT: Linsley, Peter S.
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Roberts, Chris
 ; APPLICANT: Van 't Veer, Laura
 ; APPLICANT: Van de Vijver, Marc
 ; APPLICANT: Bernard, Rene
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-188-999
 ; CURRENT APPLICATION NUMBER: US/10/172,118
 ; CURRENT FILING DATE: 2003-01-15
 ; PRIOR APPLICATION NUMBER: 60/380,770
 ; PRIOR FILING DATE: 2002-05-14
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 212
 LENGTH: 2609
 TYPE: DNA
 ORGANISM: Homo sapiens
 PUBLICATION INFORMATION:
 DATABASE ACCESSION NUMBER: AK000345
 DATABASE ENTRY DATE: 2001-06-18
 US-10-172-118-212

Query Match 61.3%; Score 18.4; DB 15; Length 2609;
 Best Local Similarity 78.6%; Pred. No. 1.9e+02; Indels 0; Gaps 0;
 Matches 22; Conservative 0; Mismatches 6;

Qy 3 GAGGATCTTGTCAAGGGCGATAAGCTGC 30
 Db 1670 GAGCATTTTCAGATGGATGGACTGC 1643

RESULT 32
 US-10-342-887-212/C
 ; Sequence 212, Application US/10342887
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yudong
 ; APPLICANT: Linsley, Peter S.
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Roberts, Christopher J.
 ; APPLICANT: Van 't Veer, Laura Johanna
 ; APPLICANT: Van de Vijver, Marc J.
 ; APPLICANT: Bernard, Rene
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-188-999
 ; CURRENT APPLICATION NUMBER: US/10/342,887
 ; CURRENT FILING DATE: 2003-01-15
 ; PRIOR APPLICATION NUMBER: 60/298,918
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: 60/380,710
 ; PRIOR FILING DATE: 2002-05-14
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 212
 LENGTH: 2609
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-342-887-212

Query Match 61.3%; Score 18.4; DB 16; Length 2609;
 Best Local Similarity 78.6%; Pred. No. 1.9e+02; Indels 0; Gaps 0;
 Matches 22; Conservative 0; Mismatches 6;

Qy 3 GAGGATCTTGTCAAGGGCGATAAGCTGC 30
 Db 1670 GAGCATTTTCAGATGGATGGACTGC 1643

RESULT 34
 US-10-425-115-34870
 ; Sequence 34870, Application US/10425115
 ; GENERAL INFORMATION:
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovacic, David K.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Plants
 ; FILE REFERENCE: 38-21(5322)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 3693326
 ; SEQ ID NO 34870
 LENGTH: 667
 TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 OTHER INFORMATION: Clone ID: MRT4577_131804C.1
 US-10-425-115-34870

Query Match 60.7%; Score 18.2; DB 18; Length 667;
 Best Local Similarity 87.0%; Pred. No. 2.2e+02; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 3;

Qy 5 GGATCTTGTCAGGCGATAAGCTGC 27
 Db 223 GGATCTTGTCAGGCGATAAGCTGC 245

RESULT 35
 US-10-425-114-22914

```

; Sequence 22914, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21153313B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 22914
LENGTH: 1523
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3592-068-C8 FLI
US-10-425-114-22914
Query Match 5 GGAATCTTGTCAAGGAGCGATAAGGC 27
Qy 5 GGAATCTTGTCAAGGAGCGATAAGGC 27
Db 876 GGAATCTTGTCAAGAAGCAGAGGGC 898
RESULT 36
US-10-425-115-34867
Sequence 34867, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
Plants
FILE REFERENCE: 38-21153222B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 3653326
SEQ ID NO 34867
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_131801C.1
US-10-425-115-34867
Query Match 5 GGAATCTTGTCAAGGAGCGATAAGGC 27
Qy 5 GGAATCTTGTCAAGGAGCGATAAGGC 27
Db 898 GGAATCTTGTCAAGAAGCAGAGGGC 920
RESULT 37
US-10-087-192-553
Sequence 553, Application US/10087192
Publication No. US20020182586A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Engelhard, Eric K.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIORITY APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIORITY APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 553
; LENGTH: 52679
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ..(52679)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-553
Query Match 4 AGGATCTGTCAAGGAGCGATAAGG 26
Qy 4 AGGATCTGTCAAGGAGCGATAAGG 26
Db 1240 AGGATCTTGCAAGGAGCGATAAGG 1262
RESULT 38
US-10-448-256-6/c
Sequence 6, Application US/10448256
Publication No. US20040043401A1
GENERAL INFORMATION:
APPLICANT: Sadelain, Michel
APPLICANT: Brentjens, Renier
APPLICANT: Maher, John
; TITLE OF INVENTION: Chimeric T Cell Receptors
; FILE REFERENCE: MSK-P-058
; CURRENT APPLICATION NUMBER: US/10/448,256
; CURRENT FILING DATE: 2003-05-28
; PRIORITY APPLICATION NUMBER: 60/383,872
; PRIOR FILING DATE: 2002-05-28
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 328
; TYPE: DNA
; ORGANISM: human
; OTHER INFORMATION: human
US-10-448-256-6
Query Match 13 TCAGGAGGGATAGGTGCG 30
Qy 13 TCAGGAGGGATAGGTGCG 30
Db 328 TCAGGAGGGATAGGTGCG 311
RESULT 39
US-09-960-352-10773
Sequence 10773, Application US/09960352
; Patent No. US20020137132A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Ningbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511,006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24

```

```

; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 10773
; LENGTH: 388
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 46-LIB34-013-Q1-E1-D6
US-09-560-352-10773

Qy      1 TCGAGGATTTCGTCAAGGGATAGG 26
Db      229 TCGAGCCTACTTGTGGAAACGAAAGG 254

```

```

RESULT 40
US-10-027-632-107348
; Sequence 107348, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 107348
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-107348

Qy      1 TCGAGGATTTCGTCAAGGGATAGG 26
Db      229 TCGAGCCTACTTGTGGAAACGAAAGG 254

```

Search completed: January 7, 2005, 14:33:04
Job time : 1291.26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 7, 2005, 10:58:39 ; Search time 1791.05 Seconds

Total number of hits satisfying chosen parameters: 65645750

Post-processing: Minimum Match length: 0
Maximum DB seq length: 2000000000

Database : EST:
 1: gb_est1: *
 2: gb_est2: *
 3: gb_htc: *
 4: gb_est3: *
 5: gb_est4: *
 6: gb_est5: *
 7: gb_est6: *
 8: gb_gss1: *
 9: gb_gss2: *

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	21	21	70.0	608	5	BQ813384	BQ813384 103003640
c	2	21	70.0	650	6	CD488918	CD488918 T16_B08 T
c	3	21	70.0	1178	5	BQ715651	BQ715651 AGENCOURT
c	4	20.8	69.3	310	9	CE517324	CE517324 tigr-gss-
c	5	20.6	68.7	664	6	CA655944	CA655944 wmo_pk0
c	6	20.4	68.0	736	9	CLJ15659	CLJ15659 104_342_1
c	7	20.4	68.0	777	5	BP700670	BP700670
c	8	20	66.7	320	6	CA122592	CA122592 SCSGSB100
c	9	20	66.7	473	8	BH051040	BH051040 RPLC-24-3
c	10	19.6	65.3	505	6	CA610501	CA610501 w1_pk011
c	11	19.6	65.3	548	8	BZ437621	BZ437621 BONP41TF
c	12	19.6	65.3	636	2	BF046684	BF046684 BP250020A
c	13	19.6	65.3	693	7	CF882870	CF882870 tric087xj
c	14	19.6	65.3	711	8	A2216274	A2216274 Sheared D
c	15	19.6	65.3	749	6	CB909216	CB909216 tric087xj
c	16	19.6	65.3	753	7	CF873769	CF873769 tric034xf
c	17	19.6	65.3	772	8	BH504182	BH504182 BCGW351F
c	18	19.6	65.3	812	8	BH687374	BH687374 BORLU53TR
c	19	19.6	65.3	823	6	CB903389	CB903389 tric034xf
c	20	19.4	64.7	318	9	CL701378	CL701378 SP_Ba007
c	21	19.4	64.7	328	6	AA486214	AA486214 ab5d07_8
c	22	19.4	64.7	390	9	CE837920	CE837920 tigr-gss-
c	23	19.4	64.7	395	6	CB773582	CB773582 AMGRNUC.N
	24	19.4	64.7	424	1	AJ706082	AJ706082

COMMENT

Contact: Charles Hauser

DCMB Box 91000

Duke University

Durham, NC 27708-1000

Tel: 919 613 8159

Fax: 919 613 8177

Email: chause@duke.edu

FEATURES

Source

/mol type="mRNA"

/strain="CC-1690 wild type mt+ 21gr"

/clone lib="C. reinhardtii CC-1690, Deflagellation

(normalized), Lambda Zap II"

/note="Vector: pBluescript II SK-, Site 1: EcoRI; Site 2:

XbaI; Deflagellation library, constructed by John Davies

and Jeffrey McDermott, combines cDNAs from CC-1690 cells

which had been re-synthesizing flagella for 15, 30 and 60

min after being deflagellated by pH shock. PolyA mRNA was

purified from each sample, pooled and cDNA synthesized.

The cDNA was directionally cloned into Lambda Zap II

(Stratagene) in the EcoRI (5') and XbaRI (3') sites.

pBluescript II SK- plasmids were excised from the lambda

ZAP clones by superinfection with EXB105 (stratagene)

phage. The library was normalized using method 4 described

in Bonaldo et al., (1996) Genome Research 6: 791-806."

ALIGNMENTS

c	25	19.4	64.7	466	9	CL702561	CL702561 Sheared D
c	26	19.4	64.7	555	8	AQ652667	AQ652667
c	27	19.4	64.7	594	9	CL700756	CL700756
c	28	19.4	64.7	613	5	BQ101535	GA_Ed000
c	29	19.4	64.7	616	1	AU274168	AU274168
c	30	19.4	64.7	620	8	BZ519836	BONQD55TF
c	31	19.4	64.7	621	7	CK086094	RG11_F05
c	32	19.4	64.7	627	9	CL703342	SP_Ba009
c	33	19.4	64.7	633	9	AG125336	Pan_trog1
c	34	19.4	64.7	634	6	CC688069	GG/BP58IV
c	35	19.4	64.7	636	8	BZ329488	hv86c01.b
c	36	19.4	64.7	668	9	CB805733	tigr-gss-
c	37	19.4	64.7	679	9	CC732980	CCYB059TV
c	38	19.4	64.7	699	9	CL697382	SP_Ba006
c	39	19.4	64.7	704	9	CL700732	SP_Ba006
c	40	19.4	64.7	710	6	CB653308	BZ696517
c	41	19.4	64.7	734	8	BZ2696517	SP_Ba008
c	42	19.4	64.7	741	4	BG281145	BZ2401647
c	43	19.4	64.7	952	1	BQ705973	AGENCO
c	44	19.4	64.7	1201	8	BZ695091	SP_Ba005
c	45	19.4	64.7	1560	3	BC022396	Homo_sapi

RESULT 1

BQ813384

LOCUS

DEFINITION

Lambda Zap II Chlamydomonas reinhardtii cdNA, mRNA sequence.

ACCESSION

BQ813384

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Chlamydomonas reinhardtii

Chlamydomonadaceae

Chlamydomonas

1 (bases 1 to 608)

REFERENCE

AUTHORS

Lefebvre, P., McDermott, J. P., Shrager, J., Silflow, C. and Stern, D.

TITLE

Analyses of the Chlamydomonas reinhardtii Genome: A Model,

Unicellular System for Analyzing Gene Function and Regulation in

Vascular Plants. Project: 1030

JOURNAL

Unpublished (2002)

COMMENT

Contact: Charles Hauser

DCMB Box 91000

Duke University

Durham, NC 27708-1000

Tel: 919 613 8159

Fax: 919 613 8177

Email: chause@duke.edu

Location/Qualifiers

1. 608

/note="Chlamydomonas reinhardtii"

/mol type="mRNA"

/strain="CC-1690 wild type mt+ 21gr"

/db_xref="taxon:3055"

/clone lib="C. reinhardtii CC-1690,

Deflagellation

(normalized), Lambda Zap II"

/note="Vector: pBluescript II SK-, Site 1: EcoRI; Site 2:

XbaI; Deflagellation library, constructed by John Davies

and Jeffrey McDermott, combines cDNAs from CC-1690 cells

which had been re-synthesizing flagella for 15, 30 and 60

min after being deflagellated by pH shock. PolyA mRNA was

purified from each sample, pooled and cDNA synthesized.

The cDNA was directionally cloned into Lambda Zap II

(Stratagene) in the EcoRI (5') and XbaRI (3') sites.

pBluescript II SK- plasmids were excised from the lambda

ZAP clones by superinfection with EXB105 (stratagene)

phage. The library was normalized using method 4 described

in Bonaldo et al., (1996) Genome Research 6: 791-806."

Query Match 70.0%; Score 21; DB 5; Length 608;
 Best Local Similarity 82.8%; Prod. No. 1.e+02;
 Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 CGAGGATCTGGTCAAGGGATAGGCTGC 30
 Db 7 CGAGGATCTGGTCAAGGGATAGGCTGC 35

RESULT 2

LOCUS CD488918 650 bp mRNA linear EST 29-AUG-2003
 DEFINITION T16_B08 Teliospore *Ustilago maydis* cDNA 5', mRNA sequence.
 ACCESSION CD488918
 VERSION EST
 SOURCE
 ORGANISM *Ustilago maydis*
Eukaryota; *Fungi*; *Basidiomycota*; *Ustilaginomycetes*; *Ustilaginaceae*; *Ustilago*.
 1 (bases 1 to 650)
 Saccadura, N.T. and Saville, B.J.
 Gene expression and EST analyses of *Ustilago maydis* germinating
 teliospores
 Fungal Genet. Biol. 40 (1), 47-64 (2003)
 22829673
 PUBMED 12948513
 COMMENT Contact: Barry J. Saville
 Saville Lab
 University of Toronto
 3359 Mississauga Road North, Mississauga, ON, L5L 1C6, Canada
 Tel: 905 569 4702
 Fax: 905 838 3792
 Email: bsaville@utoronto.ca
 Seq primer: M13 reverse primer (5' AAACAGCTATGACCATTCTCA 3').
 Location/Qualifiers 1..650
 /organism="Ustilago maydis"
 /mol_type="mRNA"
 /strain="FB2/FB2"
 /db_xref="txon:5270"
 /cell_type="teliospore"
 /dev_stage="germinating teliospore"
 /lab_host="E. coli"
 /clone_lib="Teliospore"
 /note=vector: pDNR-LIB; Site 1: SfiI; Site 2: SfiI;
 mRNA was extracted from germinating teliospores. cDNA was
 amplified by PCR and unidirectionally cloned into pDNR-LIB
 Plasmid, with the use of Clontech's Creator SMART cDNA
 Library Construction Kit."

ORIGIN

Query Match 70.0%; Score 21; DB 6; Length 650;
 Best Local Similarity 82.8%; Prod. No. 1.e+02;
 Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 CGAGGATCTGGTCAAGGGATAGGCTGC 30
 Db 623 CGAGGATCTGGTCAAGGGATAGGCTGC 595

RESULT 3

LOCUS BQ715651 1178 bp mRNA linear EST 16-JUL-2002
 DEFINITION S468577 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6308233
 5', mRNA sequence.
 ACCESSION BQ715651
 VERSION EST
 SOURCE
 ORGANISM *Mus musculus* (house mouse)
Mus musculus
Eukaryota; *Metazoa*; *Chordata*; *Caniata*; *Vertebrata*; *Euteleostomi*;

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: csgapbs@mail.nih.gov
 Tissue Procurement: Susan L. Sullivan, PhD,
 cDNA Library Preparation: ResGen, Invitrogen Corp
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at
 http://image.llnl.gov
 Plate: LIAM13727 row: a column: 02
 High quality sequence stop: 263.
 Location/Qualifiers
 1..1178
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6308233"
 /lab_host="DH10B (ophage-resistant)"
 /clone_id="NIH_MGC_129"
 /note="Organ: olfactory epithelium; Vector: PCMV-SPORT6.1;
 Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally.
 Primer: Oligo dT. Average insert size 2.2 kb. Constructed
 by ResGen, Invitrogen Corp. Note: this is a NIH-MGC
 Library."
 FEATURES
 source
 Query Match 70.0%; Score 21; DB 5; Length 1178;
 Best Local Similarity 82.8%; Prod. No. 1.e+02;
 Matches 24; Conservation 0; Mismatches 5; Index 8 0; gaps 0;
 ORIGIN

RESULT 4

LOCUS CB517324 310 bp DNA linear GSS 28-SEP-2003
 DEFINITION tigr-98s-dog-1700327420281 Dog Library *Canis familiaris* genomic,
 Genomic survey sequence.
 ACCESSION CE517324
 VERSION CE517324.1 GI:36834105
 GSS
 SOURCE *Canis familiaris* (dog)
 ORGANISM
Eukaryota; *Metazoa*; *Chordata*; *Caniata*; *Vertebrata*; *Euteleostomi*;
Mammalia; *Ustilaginomycetes*; *Ustilaginaceae*; *Ustilago*.
 1 (bases 1 to 310)
 AUTHORS Kirkness, E.F., Batra, V., Halpern, A.L., Levy, S., Remington, K.,
 Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
 Venter, J.C.
 TITLE The dog genome: survey sequencing and comparative analysis
 JOURNAL Science 301 (5641), 1898-1903 (2003)
 MEDLINE 22875432
 PUBMED 14512627
 COMMENT Contact: Kirkness, EF
 The Institute for Genomic Research, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-938-0200
 Email: ekirknes@tigr.org
 FEATURES
 source
 1..310
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /strain="Standard Poodle"

/db_xref="taxon:9615" /clone lib="Dog Library" /note="Site 1: BstX1; Libraries were prepared from peripheral Blood"	ORIGIN Query Match Score 20.8; DB 9; Length 310; Best Local Similarity 91.7%; Pred. No. 1.8e+02; Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	SOURCE ORGANISM Sorghum bicolor (sorghum) Sorghum bicolor Eukaryota; Viridiplanteae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyt; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum. 1 (bases 1 to 736)
Qy 6 GATCTTGTCAGGAGCGATAAGCTG 29 Db 1 GATCTTGTCAGGAGCGATAAGCTG 24	REFERENCE AUTHORS Budiman,M.A., Flick,B., Jones,J., Nurnberg,A., Citek,R.W., Robbins,D., Rohlfing,T., Bradford,K., Fries,J., McMenamy,J., Trani,L., Isak,A., Zimmerman,C., Lakey,N. and Bedell,J.A. Genethresher methylation filtered genomic sequences from Sorghum bicolor Unpublished (2004) Contact: Bedell, JA Orion Genomics, LLC 4041 Forest Park Ave, St. Louis, MO 63108, USA Tel: 314 615 6979 Fax: 314 615 5975 Email: jbedell@oriongenomics.com Plate: 342 row: a column: 17 Seq primer: M13/pUC Forward Class: shotgun High Quality sequence stop: 736. Location/Qualifiers 1. .736 /organism="Sorghum bicolor" /mol type="genomic DNA" /culFivar="Nx623" /db_xref="taxon:4535" /clone="10782205" /clone_lib="Sorghum methylation-filtered library (LIBID: 104)"	JOURNAL COMMENT Unpublished (2004) Contact: Bedell, JA Orion Genomics, LLC 4041 Forest Park Ave, St. Louis, MO 63108, USA Tel: 314 615 6979 Fax: 314 615 5975 Email: jbedell@oriongenomics.com Plate: 342 row: a column: 17 Seq primer: M13/pUC Forward Class: shotgun High Quality sequence stop: 736. Location/Qualifiers 1. .736 /organism="Sorghum bicolor" /mol type="genomic DNA" /culFivar="Nx623" /db_xref="taxon:4535" /clone="10782205" /clone_lib="Sorghum methylation-filtered library (LIBID: 104)"
RESULT 5 CA655944/c LOCUS CA655944 DEFINITION wlm0_pk0021.c4 wlm0 Triticum aestivum cDNA clone wlm0_pk0021.c4, end. mRNA sequence. CA655944.1 GI:25234469	FEATURES source 1. bases 1 to 664 /organism="Triticum aestivum (bread wheat)" /mol type="mRNA" /cultivar="Stephen's" /db_xref="taxon:4565" /clone="wlm0_pk0021.c4" /tissue type="leaf" /clone Lib="wlm0" /note="Vector: pBluescript SK+, Site 1: EcoRI; Site 2: XbaI; Wheat (Triticum aestivum L.) seedlings 0 hr after inoculation with Erysiphe graminis f. sp tritici"	REFERENCE AUTHORS Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z., Miao,G., Cahrah,N. and Hanafey,M.K. TITLE Dupont Wheat cDNA Sequence JOURNAL Unpublished (2002) COMMENT Contact: Scott V. Tingey E.I. DuPont de Nemours and Company 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA Tel: 302-631-2602 Fax: 302-631-2607 Email: Scott.V.Tingey@USA.dupont.com Seq primer: M13.
FEATURES source 1. .664 /organism="Triticum aestivum" /mol type="mRNA" /cultivar="Stephen's" /db_xref="taxon:4565" /clone="wlm0_pk0021.c4" /tissue type="leaf" /clone Lib="wlm0" /note="Vector: pBluescript SK+, Site 1: EcoRI; Site 2: XbaI; Wheat (Triticum aestivum L.) seedlings 0 hr after inoculation with Erysiphe graminis f. sp tritici"	ORIGIN Query Match Score 20.4; DB 9; Length 736; Best Local Similarity 80.0%; Pred. No. 3e+02; Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	REFERENCE AUTHORS Osada,S., Kitayama,A., Veno,N. and Taira,M. Title Expression analysis of Xenopus embryos JOURNAL COMMENT Unpublished (2004) Contact: Masatomi Taira Department of Biological Sciences Graduate School of Science, University of Tokyo, CREST, Japan Science and Technology Corporation, Japan 7-3-1 Horng, Bunkyo-ku, Tokyo 113-0033, Japan Tel: 81-03-5841-4434 Fax: 81-03-5841-4434
RESULT 6 CL155659/c LOCUS CL155659 DEFINITION 104_342_107822305_114_31473_017 Sorghum methylation-filtered library (LibID: 104) Sorghum_bicolor genomic clone 10782305, genomic survey sequence. ACCESSION CL155659 VERSION GI:40655575 KEYWORDS	ORIGIN Query Match Score 20.6; DB 6; Length 664; Best Local Similarity 85.2%; Pred. No. 2.4e+02; Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	REFERENCE AUTHORS Osada,S., Kitayama,A., Veno,N. and Taira,M. Title Expression analysis of Xenopus embryos JOURNAL COMMENT Unpublished (2004) Contact: Masatomi Taira Department of Biological Sciences Graduate School of Science, University of Tokyo, CREST, Japan Science and Technology Corporation, Japan 7-3-1 Horng, Bunkyo-ku, Tokyo 113-0033, Japan Tel: 81-03-5841-4434 Fax: 81-03-5841-4434
Qy 2 CGAGGATCTGTAGGAGCGATGGCT 28 Db 550 CGAGGAGCCTGTCGGACGATGGCT 524	ORIGIN Query Match Score 20.7; DB 6; Length 664; Best Local Similarity 85.2%; Pred. No. 2.4e+02; Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	REFERENCE AUTHORS Osada,S., Kitayama,A., Veno,N. and Taira,M. Title Expression analysis of Xenopus embryos JOURNAL COMMENT Unpublished (2004) Contact: Masatomi Taira Department of Biological Sciences Graduate School of Science, University of Tokyo, CREST, Japan Science and Technology Corporation, Japan 7-3-1 Horng, Bunkyo-ku, Tokyo 113-0033, Japan Tel: 81-03-5841-4434 Fax: 81-03-5841-4434

FEATURES	Source	Email: m_taira@biol.s.u-tokyo.ac.jp, URL: http://www.shigen.nig.ac.jp/nbfp/xenopus/est/ . 1 .777 /organism="Xenopus laevis" /mol_type="mRNA" /db_Xref="taxon:8355" /clone="XL450h05ex" /tissue type="anterior neuroectoderm" /dev_stage="late gastrula (stage 12.5)" /dev_stage="Osada Taira anterior neuroectoderm (ANE)" /clone_lib="Osada Taira anterior neuroectoderm (ANE)" pcsl05 cdNA library"	Query Match 68.0%; Score 20.4; DB 5; Length 777; Best Local Similarity 80.0%; Pred. No. 4.1e+02; Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
ORIGIN		RESULT 9 LOCUS BH051040/C DEFINITION RPCI-24-397C22-TV RPCI-24 Mus musculus genomic survey sequence. VERSION BH051040.1 GI:14842034 ACCESSION BH051040 KEYWORDS GSS . ORGANISM Mus musculus (house mouse) REFERENCE 1 (bases 1 to 473) AUTHORS Zhao,S., Nieman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M., Tsengye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E., Russell,D., de Jong,P. and Fraser,C.M. TITLE Mouse BAC End Sequences from Library RPCI-24 JOURNAL Other GSS: RPCI-24-397C22-TJ COMMENT Contact: Shaving Zhao Department of Bukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: shao@tigr.org Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.chor.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/orderingframe.htm). BAC end page: http://www.tigr.org/tb/bac_ends/mouse/bac_end_intro.html Plate: 397 row: C column: 22 Seq primer: T7 Class: BAC ends. FEATURES source 1. 473 /organism="Mus musculus" /mol type="genomic DNA" /strain="C57BL/6J" /db_xref="taxon:10930" /clone="RPCI-24-397C22" /sex="Male" /cell_type="Spleen/Brain" /clone lib="RPCI-24" /note="Vector: pTARBAC1; Site 1: BamH1; Site 2: BamH1. The library was cloned in the pTARBAC1 cloning vector at the BamH1 sites using MboI partially digested male C57BL/6J DNA."	ORIGIN
FEATURES	Source	Query Match 66.7%; Score 20; DB 8; Length 473; Best Local Similarity 82.1%; Pred. No. 4.3e+02; Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0; /note="Organ: ScaL Bark from adult plants; Vector: pSBI; Site 1: Sali; Site 2: NotI; An unidirectional cDNA library generated from [ScaL Bark from adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a Sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://auest.lad.ic.unicamp.br/public"	Query Match 66.7%; Score 20; DB 8; Length 473; Best Local Similarity 82.1%; Pred. No. 4.3e+02; Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
FEATURES	Source	RESULT 10 LOCUS CA610501 DEFINITION wr1.pk0117.c5 mRNA linear EST 21-NOV-2002 /db_xref="taxon:1547" /lab_host="SCS08H04" /clone lib="SBI" /note="Organ: ScaL Bark from adult plants; Vector: pSBI; Site 1: Sali; Site 2: NotI; An unidirectional cDNA library generated from [ScaL Bark from adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a Sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://auest.lad.ic.unicamp.br/public"	Query Match 66.7%; Score 20; DB 8; Length 473; Best Local Similarity 82.1%; Pred. No. 4.3e+02; Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
ORIGIN		Qy 3 GAGGATCTGTAGGAGGATAGGTGC 30 Db 225 GAGCATCTCTGAGGAAACGCTGC 252	Qy 3 GAGGATCTGTAGGAGGATAGGTGC 30 Db 343 GAGATCTCTGAGGCTATGCC 316

end, mRNA sequence.

ACCESSION CA610501
VERSION CA610501.1
SOURCE EST.
ORGANISM Triticum aestivum (bread wheat)
REFERENCE Tingey, S.V., Powell, W., Wolters, P., Dolan, M., Hainey, C., Yuan, Z., Miao, G., Caraher, N., and Hanafey, M.K.
AUTHORS E.I. DuPont de Nemours and Company
TITLE Dupont Wheat cDNA Sequence
JOURNAL Unpublished (2002)
COMMENT Contact: Scott V. Tingey
Crop Genetics
E.I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq Primer: M13.

FEATURES source
LOCUS BONPJ41TF EST.
DEFINITION BONPJ41TF EO_1_6_2_KB_tot
Organism: *Triticum aestivum*
/mol type="mRNA"
/db_xref="taxon:3712"
/clone="BONPJ41"
/clone_lid="BO_1_6_2_KB_tot"
/note="Vector: phOSI-, Site1: BstXI, 1.6-2 kb sheared total DNA inserted into phOSI using BstXI linkers"

Query Match 65.3%; Score 19.6; DB 8;
Best Local Similarity 84.6%; Pred. No. 6.7e+03;
Matches 22; Conservative 0; Mismatches 4;
Indels 0; Gaps 0;

ORIGIN

Query Match 65.3%; Score 19.6; DB 8;
Best Local Similarity 84.6%; Pred. No. 6.7e+03;
Matches 22; Conservative 0; Mismatches 4;
Indels 0; Gaps 0;

Qy 4 AGGATCTGTAGGGATAGGCTG 29
Db 89 AGGATCTGTAGGGATAGGCTG 64

RESULT 12

BF046684 LOCUS BP046684 636 bp mRNA linear EST 10-OCT-2000
DEFINITION BP250020A10B3 Soares normalized bovine placenta Bos taurus cDNA clone BP250020A10B3 5', mRNA sequence.
ACCESSION BF046684
VERSION BF046684.1 GI:10763728
KEYWORDS BST.
SOURCE Bos taurus (cow)
ORGANISM Bovaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 636)
AUTHORS Lewin, H.A.; Soares, M.B.; Rebeiz, M.; Pardinas, J.; Liu, L. and Larson, J.H.
TITLE Bovine ESTs
JOURNAL Unpublished (2000)
COMMENT Contact: Lewin, H. A.
W. M. Keck Center for Comparative and Functional Genomics
University of Illinois at Urbana-Champaign
340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL
61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Funding for cattle EST sequencing was provided by the USDA National Research Initiative, Animal Genome Resource Grant AG 99-3205-8534 to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED from Washington University Genome Center. Vector Trimmer: G; Cross match from Washington University Genome Center PHRAP suite. This sequence is vector free and at least 200 bp in length.
PCR PRIMERS FORWARD: TAATAGCACTACTATAGGG
BACKWARD: ATTAAACCCTCACTTAAG
Insert Length: 636 Std Error: 0.00
Plate: BP250020A10 row: B column: 3
Seq primer: AGCGGATPACAAATTCAACACGGA
High quality sequence stop: 636.
FEATURES source
ORGANISM Brassica oleracea
DEFINITION BONPJ41TF EO_1_6_2_KB_tot
Organism: *Brassica oleracea*; Streptophytina; Embryophyta; Tracheophyta; Spermatophytina; Magnoliophyta; eudicots; core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
/note="1 (bases 1 to 548)"
/clone="BONPJ41"
/clone_lid="Soare normalized bovine placenta"
/note="Organ: Placenta; vector: pRT3Pac; Site: 1; ECORI; Site 2; Not I; The cDNA library was contributed by the Soares laboratory and it was constructed and normalized as described by Bonaldo, M.P., Lennon, G. and Soares, M.B. (1996), Genome Research 6 (9): 791-806."

RESULT 11

BZ437621/c LOCUS BZ437621 548 bp DNA linear GSS 13-DEC-2002
DEFINITION BONPJ41TF EO_1_6_2_KB_tot
Organism: *Brassica oleracea*; Streptophytina; Embryophyta; Tracheophyta; Spermatophytina; Magnoliophyta; eudicots; core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
/note="Whole genome shotgun sequencing of *Brassica oleracea*"
/clone="BONPJ41"
/clone_lid="Soare normalized bovine placenta"
/note="Organ: Placenta; vector: pRT3Pac; Site: 1; ECORI; Site 2; Not I; The cDNA library was contributed by the Soares laboratory and it was constructed and normalized as described by Bonaldo, M.P., Lennon, G. and Soares, M.B. (1996), Genome Research 6 (9): 791-806."

REFERENCE Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
AUTHORS Tom, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
TITLE Unpublished (2001)
JOURNAL Other GSSs: BONPJ41TR
COMMENT Contact: Chris Town
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-88-3523
Fax: 301-88-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TP
Class: sheared ends.
FEATURES source
ORGANISM Brassica oleracea
DEFINITION BONPJ41TR
Organism: *Brassica oleracea*
/mol type="mRNA"
/db_xref="taxon:9913"
/clone="BP250020A10B3"
/sex=female"
/lab host="DH10B"
/note="Soare normalized bovine placenta"
/clone_lid="Soare normalized bovine placenta"
Site: 1: ECORI;
Site 2: Not I; The cDNA library was contributed by the Soares laboratory and it was constructed and normalized as described by Bonaldo, M.P., Lennon, G. and Soares, M.B. (1996), Genome Research 6 (9): 791-806."

Query Match 65.3%; Score 19.6; DB 2; Length 636;
Best Local Similarity 84.6%; Pred. No. 6.8e+02; Malek,J., Fujii,C.,
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
TITLE
JOURNAL
COMMENT

REFERENCE
AUTHORS El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,B., Malek,J., Fujii,C.,
Gerrard,C., Leech,V., de Jong,P., Ullu,E., Melville,S.,
Doneilson,J., Fraser,C. and Adams,M.
DEFINITION Determination of clone end sequences from *Trypanosoma brucei* GUTat
10.1 Sheared DNA library
Unpublished DNA
Db 141 GCATCTGGAGGTAGTGTAGCTGC 166

Trypanosoma 1 (bases 1 to 711)
El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,B., Malek,J., Fujii,C.,
Gerrard,C., Leech,V., de Jong,P., Ullu,E., Melville,S.,
Doneilson,J., Fraser,C. and Adams,M.
DEFINITION Determination of clone end sequences from *Trypanosoma brucei* GUTat
10.1 Sheared DNA library
Unpublished DNA
Db 141 GCATCTGGAGGTAGTGTAGCTGC 166

Other GSS: Sheared DNA-57B2-TF
Contact: Najib M. El-Sayed
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: nel-sayed@tigr.org
Clones are derived from the *Trypanosoma brucei* GUTat 10.1 sheared
DNA library constructed at TIGR. Clones will be available for
distribution through Research Genetics, Alabama, USA. Sheared DNA
end sequences search page: <http://www.tigr.org/tgb/mdb/>.
Seq primer: M13-Reverse
Class: shotgun
Location/Qualifiers

FEATURES source
1. .711
/organism="Trypanosoma brucei"
/mol_type="Genomic DNA"
/strain="TRBU927/4 GUTat 10.1"
/note="Vector: pUC18; Site 1: Smal; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TRBU927/4 GUTat 10.1) was mechanically
sheared to give a tight size distribution (approx 2 kb).
The v + i method used for the library construction is
described in detail in Smith, H.O. and Venter, J.C.
(Making small insert libraries for whole genome shotgun
sequencing projects. In Genome Sequencing: A Practical
Approach, eds. M. Vaudin and B. Barrell, Oxford University
Press, 1999)."

ORIGIN
Query Match 65.3%; Score 19.6; DB 8; Length 711;
Best Local Similarity 84.6%; Pred. No. 6.9e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
DEFINITION tric087x104 T. reesei mycelial culture, Version 3 april Hypocrea
jecorina CDNA clone tric087x104, mRNA sequence.

ACCESSION CB909216/c
VERSION CB909216.1
KEYWORDS EST.
ORGANISM Hypocrea jecorina (anamorph: Trichoderma reesei)
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocreae.

RESULT 15
CB909216/c
LOCUS CB909216
DEFINITION tric087x104 T. reesei mycelial culture, Version 3 april Hypocrea
jecorina CDNA clone tric087x104, mRNA sequence.

Qy 1 TCGGGATCTGGAGGGATAGG 26
Db 165 TCGGGATCTGGAGGGATAGG 140

ORIGIN
Query Match 65.3%; Score 19.6; DB 7; Length 693;
Best Local Similarity 84.6%; Pred. No 6.9e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
DEFINITION Sheared DNA-57B2-TR Sheared DNA *Trypanosoma brucei* genomic clone
Db 165 TCGGGATCTGGAGGGATAGG 140

RESULT 14
AZ216274/c
LOCUS AZ216274
DEFINITION Sheared DNA-57B2-TR Sheared DNA *Trypanosoma brucei* genomic survey sequence.
ACCESSION AZ216274
VERSION AZ216274.1
KEYWORDS GSS
SOURCE Trypanosoma brucei
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatida;

RESULT 14
AZ216274/c
LOCUS AZ216274
DEFINITION Sheared DNA-57B2-TR Sheared DNA *Trypanosoma brucei* genomic survey sequence.
ACCESSION AZ216274
VERSION AZ216274.1
KEYWORDS GSS
SOURCE Trypanosoma brucei
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatida;

REFERENCE
AUTHORS Foreman,P.K., Brown,D.E., Dankmeyer,L., Dean,R., Diener,S.,
Dunn-Coleman,N.S., Goedgebeur,F., Houfek,T.D., England,G.J.,
Kelle, A.S., Neerman,H.J., Mitchell,T., Mitchison,C.,
Olivares,H.A., Tenissen,P.J., Yeo,J. and Ward,M.
DEFINITION Transcriptional regulation of biomass-degrading enzymes in the
filamentous fungus *Trichoderma reesei*
JOURNAL J. Biol. Chem. 278 (34), 31988-31997 (2003)
MEDLINE 22803314

PUTMED	12788920	COMMENT	Contact: Pamela K. Foreman Generacor Int'l. Road, Palo Alto, CA 94304, USA Tel: (650) 846-7635 Fax: (650) 621-7817 Email: pforeman@gencor.com Seq primer: LT_F1 primer Location/Qualifiers	source	Query Match 65.3%; Score 19.6; DB 7; Length 753; Best Local Similarity 84.6%; pred. No. 7e+02; Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
FEATURES					
		ORIGIN	1. .749 /organism="Hypocrea jecorina" /mol_type="mRNA" /strain="Qn6" /db_xref="Taxon:51451" /clone="tric087xj04" /dev_stage="mycelia" /clone_lib="T. reesei mycelial culture, Version 3 April" /note="Vector: PREP3; Site_1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."	RESULT 17 BH504182 DEFINITION BOGQW35TR BOQGQW35 genomic clone BOQW35, genomic survey sequence. ACCESSION BH504182 VERSION BH504182.1 GI:17712279 KEYWORDS GSS. SOURCE Brassica oleracea ORGANISM Brassica oleracea Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.	
		ORIGIN	2. .749 /organism="Hypocrea jecorina" /mol_type="mRNA" /strain="Qn6" /db_xref="Taxon:51451" /clone="tric087xj04" /dev_stage="mycelia" /clone_lib="T. reesei mycelial culture, Version 3 April" /note="Vector: PREP3; Site_1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."	RESULT 17 BH504182 DEFINITION BOGQW35TR BOQGQW35 genomic clone BOQW35, genomic survey sequence. ACCESSION BH504182 VERSION BH504182.1 GI:17712279 KEYWORDS GSS. SOURCE Brassica oleracea ORGANISM Brassica oleracea Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.	
		ORIGIN	3. .749 /organism="Hypocrea jecorina" /mol_type="mRNA" /strain="Qn6" /db_xref="Taxon:51451" /clone="tric087xj04" /dev_stage="mycelia" /clone_lib="T. reesei mycelial culture, Version 3 April" /note="Vector: PREP3; Site_1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."	RESULT 17 BH504182 DEFINITION BOGQW35TR BOQGQW35 genomic clone BOQW35, genomic survey sequence. ACCESSION BH504182 VERSION BH504182.1 GI:17712279 KEYWORDS GSS. SOURCE Brassica oleracea ORGANISM Brassica oleracea Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.	
FEATURES		ORIGIN	4. .749 /organism="Hypocrea jecorina" /mol_type="mRNA" /strain="Qn6" /db_xref="Taxon:51451" /clone="tric087xj04" /dev_stage="mycelia" /clone_lib="T. reesei mycelial culture, Version 3 April" /note="Vector: PREP3; Site_1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."	RESULT 17 BH504182 DEFINITION BOGQW35TR BOQGQW35 genomic clone BOQW35, genomic survey sequence. ACCESSION BH504182 VERSION BH504182.1 GI:17712279 KEYWORDS GSS. SOURCE Brassica oleracea ORGANISM Brassica oleracea Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.	
PUTMED	CF973769/c	COMMENT			
FEATURES		ORIGIN	1. .753 /organism="Hypocrea jecorina" /mol_type="mRNA" /strain="Qn6" /db_xref="Taxon:51451" /clone="tric034xf13" /dev_stage="mycelia" /clone_lib="T. reesei mycelial culture, Version 6 October 2003" /note="Vector: PREP3; Site_1: Not I/Sal I; mRNA sequence." /clone="tric034xf13" /note="Vector: PREP3; Site_1: Not I/Sal I; mRNA sequence." /clone="tric034xf13"	RESULT 16 CF973769/c LOCUS CF973769 DEFINITION Hypocrea jecorina cDNA clone tric034xf13, mRNA sequence. ACCESSION CF973769 VERSION CF973769.1 KEYWORDS EST SOURCE Hypocrea jecorina (amamorph: Trichoderma reesei) ORGANISM Hypocrea jecorina Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocreae. REFERENCE 1. (bases 1 to 753) AUTHORS Diener, S.E., Dunn-Coleman,N., Foreman,P., Houfek,T.D., Teunissen,P.J.M., van Solingen,P., Mitchell,T.K., Ward,M. and Dean,R.A. TITLE Characterization of the protein processing and secretion pathways in a comprehensive set of expressed sequence tags from Trichoderma reesei JOURNAL FEMS Microbiol. Lett. 230 (2), 275-282 (2004) COMMENT Contact: Ralph A. Dean Fungal Genomics Laboratory North Carolina State University Campus Box 7251, Raleigh, NC 27695, USA Tel: 919-515-0020 Fax: 919-513-0024 Email: ralph.dean@ncsu.edu Seq primer: LT_F1 primer Location/Qualifiers	Query Match 65.3%; Score 19.6; DB 8; Length 772; Best Local Similarity 84.6%; pred. No. 7e+02; Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
FEATURES		ORIGIN	2. .753 /organism="Hypocrea jecorina" /mol_type="mRNA" /strain="Qn6" /db_xref="Taxon:51451" /clone="tric034xf13" /dev_stage="mycelia" /clone_lib="T. reesei mycelial culture, Version 6 October 2003" /note="Vector: PREP3; Site_1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."	RESULT 18 BH687374/c LOCUS BH687374 DEFINITION BOHDX53TR BO_2_3 KB Brassica oleracea genomic clone BOHUX53, genomic survey Sequence. ACCESSION BH687374 VERSION BH687374.1 GI:18757811 KEYWORDS GSS. SOURCE Brassica oleracea ORGANISM Brassica oleracea Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica. REFERENCE 1. (bases 1 to 812) AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.	Query Match 65.3%; Score 19.6; DB 8; Length 772; Best Local Similarity 84.6%; pred. No. 7e+02; Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
FEATURES		ORIGIN	3. .753 /organism="Hypocrea jecorina" /mol_type="mRNA" /strain="Qn6" /db_xref="Taxon:51451" /clone="tric034xf13" /dev_stage="mycelia" /clone_lib="T. reesei mycelial culture, Version 6 October 2003" /note="Vector: PREP3; Site_1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."	RESULT 18 BH687374/c LOCUS BH687374 DEFINITION BOHDX53TR BO_2_3 KB Brassica oleracea genomic clone BOHUX53, genomic survey Sequence. ACCESSION BH687374 VERSION BH687374.1 GI:18757811 KEYWORDS GSS. SOURCE Brassica oleracea ORGANISM Brassica oleracea Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica. REFERENCE 1. (bases 1 to 812) AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.	Query Match 65.3%; Score 19.6; DB 8; Length 772; Best Local Similarity 84.6%; pred. No. 7e+02; Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
FEATURES		ORIGIN	4. .753 /organism="Hypocrea jecorina" /mol_type="mRNA" /strain="Qn6" /db_xref="Taxon:51451" /clone="tric034xf13" /dev_stage="mycelia" /clone_lib="T. reesei mycelial culture, Version 6 October 2003" /note="Vector: PREP3; Site_1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."	RESULT 18 BH687374/c LOCUS BH687374 DEFINITION BOHDX53TR BO_2_3 KB Brassica oleracea genomic clone BOHUX53, genomic survey Sequence. ACCESSION BH687374 VERSION BH687374.1 GI:18757811 KEYWORDS GSS. SOURCE Brassica oleracea ORGANISM Brassica oleracea Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.	Query Match 65.3%; Score 19.6; DB 8; Length 772; Best Local Similarity 84.6%; pred. No. 7e+02; Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

TITLE	Whole genome shotgun sequencing of <i>Brassica oleracea</i>	ORIGIN	and Nitrogen sources and concentrations."
JOURNAL	Unpublished (2001)		
COMMENT	Other_GSSs_BOHUX53TF Contact: Chris Town 9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-832-3523 Fax: 301-832-0208 Email: cdrown@tigr.org DNA is from doubled haploid provided by Tom Osborn. Seq primer: TR Class: sheared ends.		
FEATURES	Location/Qualifiers		
source	1. .812 /organism="Brassica oleracea" /mol_type="genomic DNA" /strain="TIGR000DH3" /db_xref="taxon:3712" /clone="BOHUX53" /cln_lib="BO_2_3 KB" /note="Vector: PHOSI; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOSI using BstXI linkers"		
ORIGIN			
Query Match	Score 19.6; DB 8; Length 812;		
LOCUS	CB903389	823 bp mRNA	EST 02-JUL-2003
DEFINITION	tric03xf13 T.reesei mycelial culture, Version 3 april Hypocreae		
KEYWORDS	secorina cDNA clone tric034xf13, mRNA sequence.		
ACCESSION	CB903389		
VERSION	EST		
SOURCE	4 AGGATCTGTCAAGGCGTAGGCTG 29		
ORGANISM	394 AGGATCTTCATGAAAGAAGGTG 369		
RESULT 19			
LOCUS	CB903389/c		
DEFINITION	tric03xf13 T.reesei mycelial culture, Version 3 april Hypocreae		
KEYWORDS	secorina cDNA clone tric034xf13, mRNA sequence.		
ACCESSION	CB903389		
VERSION	EST		
SOURCE	4 AGGATCTGTCAAGGCGTAGGCTG 29		
ORGANISM	394 AGGATCTTCATGAAAGAAGGTG 369		
RESULT 20			
LOCUS	CL701378	318 bp	DNA
DEFINITION	SP_Ba0072B17_r SP_Ba Sorghum propinquum Genomic clone		
ACCESSION	CL701378		
VERSION	GSS		
KEYWORDS	sorghum propinquum		
SOURCE	ORGANISM		
ORGANISM	Sorghum propinquum		
Bakteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyt; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.			
REFERENCE	1 (bases 1 to 318)		
AUTHORS	Kim.H., Yu,Y., Wissotski,M., Yost,D., Rao,K., Kudrna,D., Muller,C., Soderlund,C., Bowers,J.E., Paterson,A.H. and Wing,R.		
TITLE	Physical mapping of the sorghum genome		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: Rod A. Wing Arizona Genomics Institute University of Arizona Forbes Building Room 303, Tucson, AZ 85721-0036, USA Tel: 520 626 9595 Fax: 520 621 1259 Email: http://genome.arizona.edu PCR Primers		
FEATURES	source		
source	1. .318 /organism="Sorghum propinquum" /mol_type="genomic DNA" /db_xref="taxon:132711" /clone="SP_Ba0072B17" /clone lib="SP_Ba" /note="Vector: pBelobAC1; Site 1: HindIII; Site 2: HindIII; Paterson lab BAC library (HindIII)"		
ORIGIN			
Query Match	Score 19.4; DB 9; Length 318;		
LOCUS	AA486214	328 bp mRNA	EST 06-MAR-1998
DEFINITION	ab35607_s1 Stratagene HeLa cell s3 93716 Homo sapiens cDNA clone IMAGE:8422797 3' similar to gb:S71381 PROTEASOME BETA CHAIN (HUMAN); mRNA sequence.		
ACCESSION	AA486214		
VERSION	2 CGGGATCTGTCAAGGCGTAGGCTG 30		
SOURCE	Db 108 CGACTATCTTCATGACCAATTGGCTGC 136		
RESULT 21			
LOCUS	AA486214	328 bp mRNA	EST 06-MAR-1998
DEFINITION	ab35607_s1 Stratagene HeLa cell s3 93716 Homo sapiens cDNA clone IMAGE:8422797 3' similar to gb:S71381 PROTEASOME BETA CHAIN (HUMAN); mRNA sequence.		
ACCESSION	AA486214		
VERSION	2 CGGGATCTGTCAAGGCGTAGGCTG 30		
SOURCE	Db 108 CGACTATCTTCATGACCAATTGGCTGC 136		
FEATURES	Location/Qualifiers		
source	1. .823 /organism="Hypocrease jecorina" /mol_type="mRNA" /strin="OMg6" /db_xref="taxon:51453" /clone="tric034xf13" /dev_stge="mycelia" /clone lib="T.reesei mycelial culture, Version 3 april" /note="Vector: PR3PY; Site 1: Not I; Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon		
COMMENT	955 Page Mill Road, Palo Alto, CA 94304, USA Tel: (650) 846-7635 Fax: (650) 621-7817 Email: Pforeran@gennencor.com Seq primer: LR-F1 primer.		
FEATURES	Location/Qualifiers		
source	1. .823 /organism="Hypocrease jecorina" /mol_type="mRNA" /strin="OMg6" /db_xref="taxon:51453" /clone="tric034xf13" /dev_stge="mycelia" /clone lib="T.reesei mycelial culture, Version 3 april" /note="Vector: PR3PY; Site 1: Not I; Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon		

KEYWORDS	EST.	Homo sapiens (human)
SOURCE	ORGANISM	Homo sapiens
REFERENCE	Mammalia; Eutheria; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;	
AUTHORS	1 (bases 1 to 424); Depita,C., Tombolan,L., Kronnie,G., Romualdi,C., Vitulo,N., Basso,G. and Lanfranchi,G.	
TITLE	A leukemia-enriched cDNA microarray platform identified new transcripts with relevance to the biology of leukemias	
JOURNAL	Unpublished (2004)	
COMMENT	Contact : Depita C Biology and CRIBI University of Padova Via U. Bassi, 58/B, 35131, ITALY.	
FEATURES	source	Location/Qualifiers 1. .424 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clones="BLPD01229" /issue_type="bone marrow" /clone_lib="BLPD01" /note="caucasian"
ORIGIN	Query Match 64.7%; Score 19.4; DB 1; Length 424; Best Local Similarity 79.3%; Pred. No. 7.9e-02; Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
FEATURES	source	Location/Qualifiers 1. .424 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clones="BLPD01229" /issue_type="bone marrow" /clone_lib="BLPD01" /note="caucasian"
ORIGIN	Query Match 64.7%; Score 19.4; DB 1; Length 424; Best Local Similarity 79.3%; Pred. No. 7.9e-02; Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
FEATURES	source	Location/Qualifiers 1. .424 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clones="BLPD01229" /issue_type="bone marrow" /clone_lib="BLPD01" /note="caucasian"
ORIGIN	Query Match 64.7%; Score 19.4; DB 1; Length 424; Best Local Similarity 79.3%; Pred. No. 8e+02; Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
FEATURES	source	Location/Qualifiers 1. .466 /organism="Sorghum propinquum" /mol_type="genomic DNA" /db_xref="taxon:132711" /clone_lib="SP_Ba0085N23" /note="PCR primers"
ORIGIN	Query Match 64.7%; Score 19.4; DB 8; Best Local Similarity 79.3%; Pred. No. 8.2e+02; Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	

Db	30	TGGAGGATCCTTACGGCGATGGGTG	2		COMMENT	Contact: Wing RA Clemson University Genomics Institute Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA
RESULT 27					LOCUS	CL700756
CL700756	CL700756	594 bp	DNA	linear	GSS	13-JUL-2004
LOCUS	CL700756	594 bp	DNA	linear	GSS	13-JUL-2004
DEFINITION	SP_Ba0065A04_f SP_Ba0065A04_5'	Ba Sorghum propinquum genomic clone			DEFINITION	SP_Ba0065A04_f, genomic survey sequence.
ACCESSION	CL700756				ACCESSION	BA0065008f
VERSION	CL700756.1				VERSION	GI:50267031
KEYWORDS					KEYWORDS	Sorghum propinquum
SOURCE					SOURCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD Clade; Panicoideae; Andropogoneae; Sorghum.
ORGANISM					ORGANISM	1. (bases 1 to 594)
COMMENT					COMMENT	Kim.H., Yu.Y., Wissotski,M., Yst,D., Stum,D., Rao,K., Kudrna,D., Muller,C., Soderlund,C., Bowers,J.E., Paterson,A.H. and Wing,R. Physical mapping of the sorghum genome unpublished (2004).
REFERENCE					REFERENCE	
AUTHORS					AUTHORS	
TITLE					TITLE	
JOURNAL					JOURNAL	
FEATURES					FEATURES	
Source					Source	
1.	.594				1.	.594
/organism="Sorghum propinquum"					/organism="Sorghum propinquum"	
/mol type="genomic DNA"					/mol type="genomic DNA"	
/db_xref="taxon:132211"					/db_xref="taxon:132211"	
/clone="SP_Ba0065A04"					/clone="SP_Ba"	
/clone lib="SP_Ba"					/clone lib="Vector: BeloBAC1"	
/note="Vector: BeloBAC1; Site_1: HindIII; Site_2: HindIII; Paterson lab BAC library (HindIII)"					/note="Vector: BeloBAC1; Site_1: HindIII; Site_2: HindIII; Paterson lab BAC library (HindIII)"	
CLASS: BAC ends.					CLASS: BAC ends.	
Location/Qualifiers					Location/Qualifiers	
ORIGIN					ORIGIN	
COMMENT					COMMENT	
REFERENCE					REFERENCE	
AUTHORS					AUTHORS	
TITLE					TITLE	
JOURNAL					JOURNAL	
FEATURES					FEATURES	
Source					Source	
1.	.594				1.	.594
/organism="Sorghum propinquum"					/organism="Sorghum propinquum"	
/mol type="genomic DNA"					/mol type="genomic DNA"	
/db_xref="taxon:132211"					/db_xref="taxon:132211"	
/clone="SP_Ba0065A04"					/clone="SP_Ba"	
/clone lib="SP_Ba"					/clone lib="Vector: BeloBAC1"	
/note="Vector: BeloBAC1; Site_1: HindIII; Site_2: HindIII; Paterson lab BAC library (HindIII)"					/note="Vector: BeloBAC1; Site_1: HindIII; Site_2: HindIII; Paterson lab BAC library (HindIII)"	
CLASS: BAC ends.					CLASS: BAC ends.	
Location/Qualifiers					Location/Qualifiers	
ORIGIN					ORIGIN	
COMMENT					COMMENT	
REFERENCE					REFERENCE	
AUTHORS					AUTHORS	
TITLE					TITLE	
JOURNAL					JOURNAL	
FEATURES					FEATURES	
Source					Source	
1.	.594				1.	.594
/organism="Sorghum propinquum"					/organism="Sorghum propinquum"	
/mol type="genomic DNA"					/mol type="genomic DNA"	
/db_xref="taxon:132211"					/db_xref="taxon:132211"	
/clone="SP_Ba0065A04"					/clone="SP_Ba"	
/clone lib="SP_Ba"					/clone lib="Vector: BeloBAC1"	
/note="Vector: BeloBAC1; Site_1: HindIII; Site_2: HindIII; Paterson lab BAC library (HindIII)"					/note="Vector: BeloBAC1; Site_1: HindIII; Site_2: HindIII; Paterson lab BAC library (HindIII)"	
CLASS: BAC ends.					CLASS: BAC ends.	
Location/Qualifiers					Location/Qualifiers	
ORIGIN					ORIGIN	
COMMENT					COMMENT	
REFERENCE					REFERENCE	
AUTHORS					AUTHORS	
TITLE					TITLE	
JOURNAL					JOURNAL	

Qy	1 TCGAGGATCTTGTCAAGGGATAGGCTG	29	MI 48824, USA
Db	515 TCGAGGATCTTGTCAAGGGATAGGCTG	487	Tel: 517 353 0890 Fax: 517 355 5191 x431
RESULT 30			Email: grumet@msu.edu
LOCUS	BZ519836	620 bp	Plate: RG11 Row: F column: 05.
DEFINITION	BOMQD55TF BO_2_3_KB Brassica oleracea genomic clone BOMQD55,		Location/Qualifiers
ACCESSION	BZ519836		1..621
VERSION	BZ519836.1		/organism="Cucumis sativus"
KEYWORDS			/mol_type="mRNA"
SOURCE	Brassica oleracea		/strand="Straight 8"
ORGANISM	Bukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina;		/db_xref="Taxon:3659"
	Spermatophyta; Magnoliophytina; eudicotyledons; core eudicots;		/sex="monocious"
	robdids; euroids II; Brassicales; Brassicaceae; Brassicas.		/clone lib="Cucumber leaf"
REFERENCE			/note="Vector: PAD-GAL4; Site_1: EcoRI; Site_2: XbaI"
AUTHORS	Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.		
TITLE	Whole genome shotgun sequencing of Brassica oleracea		
JOURNAL	Unpublished (2001)		
COMMENT	Other_GSSs: BOMQD55TR		
TIGR	9712 Medical Center Drive, Rockville, MD 20850, USA.		
	Tel: 301-888-3523		RESULT 32
	Fax: 301-888-0208		CL703342
	Email: cdtown@tigr.org		LOCUS
	DNA is from a doubled haploid provided by Tom Osborn.		SP_Ba0095F13_r SP_Ba Sorghum propinquum genomic clone
	seq primer: TF		SP_Ba0095F13_3', genomic survey sequence.
	Class: sheared ends.		DEFINITION
FEATURES	Source		CL703342
	1..620		ACCESSION
	/organism="Brassica oleracea"		CL703342.1 GI:50269617
	/mol_type="genomic DNA"		GSS.
	/strain="TO_0000013"		Sorghum propinquum
	/db_xref="taxon:3712"		Eukaryota; Viridiplantae; Streptophytina; Tracheophytina;
	/clone="BOMQD55"		Spermatophyta; Magnoliophytina; Liliopsida; Poales; Poaceae; PACCAD
	/clone lib="BO_2_3_KB"		clade; Panicoideae; Andropogoneae; Sorghum.
	/note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"		1 (bases 1 to 627)
ORIGIN			REFERENCE
	Query Match 64.7%; Score 19.4; DB 8; Length 620;		AUTHORS
	Best Local Similarity 79.3%; Pred. No. 8.e+02;		Kim,H., Yu,Y., Wissotski,M., Yost,D., Stum,D., Rao,K., Kudrna,D.,
	Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;		Muller,C., Soderblom,C., Yersin,J.E., Peterson,A.H. and Wing,R.
Qy	2 CGAGATCTTGTCAAGGGATAGGCTG	30	TITLE
Db	590 CCAGCTTCCTGTGAGGAACGATGCCCTGC	562	JOURNAL
			Physical mapping of the sorghum genome
			Unpublished (2004)
			Contact: Rod A. Wing
			Arizona Genomics Institute
			Forbes Building Room 303, Tucson, AZ 85721-0036, USA
RESULT 31			COMMENT
LOCUS	CK086094	621 bp	FORWARD: TAA TAC GAC TCA CTA TAG GG
DEFINITION	FG11_F05 Cucumber leaf Cucumis sativus cDNA, mRNA sequence.		BACKWARD: CAC TCA TTA GGC ACC CCA
ACCESSION	CK086094		Plate: 0095 Row: F column: 13
VERSION	CK086094.1		Tel: 520 626 9595
KEYWORDS	EST.		Fax: 520 621 1259
ORGANISM	Cucumis sativus (cucumber)		Email: http://genome.arizona.edu
REFERENCE	Grumet,R. and McGrath,M.		PCR Primers
AUTHORS	Development of genomic tools for cucumber (Cucumis sativus L.)		FORWARD: TAA TAC GAC TCA CTA TAG GG
TITLE	Unpublished (2003)		BACKWARD: CAC TCA TTA GGC ACC CCA
JOURNAL	Rebecca Grumet		Plate: 0095 Row: F column: 13
COMMENT	Michigan State University Horticulture Department, Michigan State University, East Lansing,		Tel: 520 626 9595
Qy	2 CGAGGATCTTGTCAAGGGATAGGCTG	30	Fax: 520 621 1259

Db	173 CGCTATCTTTCAGGATCAATTGGCTGC 201	TITLE	Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
RESULT	33	JOURNAL	Consortium for Marmoset Genomics
AG125336/c	AG125336 633 bp DNA linear GSS 04-NOV-2001	COMMENT	Unpublished (2002)
LOCUS	AG125336 Pan troglodytes DNA, clone: PTB-135124.R, genomic survey sequence.	Contact	Cathy Whitelaw
DEFINITION		TIGR	
ACCESSION	AG125336	9712 Medical Center Drive, Rockville, MD 20850, USA	
VERSION	AG125336.1 GI:16654501	Te: 301-838-5843	
KEYWORDS	Pan troglodytes (chimpanzee)	Fax: 301-838-0208	
SOURCE	Pan troglodytes	Email: whitelaw@tigr.org	
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.	Seq primer: TIP	
REFERENCE	Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.	Class: sheared ends.	
AUTHORS	Totoki,Y., Watanabe,H. and Sakaki,Y.	Location/Qualifiers	
TITLE	BAC end sequences of Library PTB	1. .634	/organism="Zea mays"
JOURNAL	Unpublished		/mol type="genomic DNA"
REFERENCE	2 (bases 1. to 633) Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.		/strain="B73"
AUTHORS			/db_xref="taxon:45777"
TITLE	Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimppbs@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel: 01-45-503-9111, Fax: 01-45-503-9170)		/clone_id="ZMMBMA053720"
JOURNAL			/note="Vector: PBCSK-; Site 1: HinclII; 0.7-1.5 kb
COMMENT	Clones are derived from the chimpanzee BAC library PTB. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.		methylation filtered genomic DNA library"
PRIMERS			
SEQUENCING	Sequencing: M13RE		
LIBRARY			
VECTOR	: pk8145		
R.Site 1	: SacI		
R.Site 2	: SacI'		
FEATURES	Location/Qualifiers		
source	1..633		
DEFINITION	/organism="Pan troglodytes"		
KEYWORDS	/mol type="genomic DNA"		
ACCESSION	/db_xref="taxon:9598"		
VERSION	/clones="PTB-135124.R"		
COMMENT	/sex="male"		
ORIGIN	/cell type="lymphoblast"		
SEQUENCING	/clone_lib="PTB Chimpanzee Male BAC Library"		
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN			
SEQUENCING			
LIBRARY			
VECTOR			
R.Site 1			
R.Site 2			
FEATURES	Location/Qualifiers		
source			
DEFINITION			
KEYWORDS			
ACCESSION			
VERSION			
COMMENT			
ORIGIN	</		

end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (.x/y reads in M13mp19, .b/g reads in pUC19). The same ligation was transformed in either JM107 or DH5α."

ORIGIN

Query Match 64.7%; Score 19.4; DB 8; Length 63;
Best Local Similarity 79.3%; Pred. No. 8.4e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 2 CGGAGATCTGTCAAGGAGCATAGGTGC 30
Db 309 CGATCATTTTCAGGACCAATGGCTGC 281

RESULT 36

CE05733/c
LOCUS CE05733 668 bp DNA linear GSS 30-SEP-2003
DEFINITION tigr-qss-dog-17000331469523 Dog Library Canis familiaris genomic, genomic survey sequence.

ACCESSION CE05733
VERSION CE05733.1 GI:371466555
KEYWORDS GSS
SOURCE Canis familiaris (dog)

ORGANISM Eukaryota; Metacoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Buteraria; Carnivora; Fissipedia; Canidae; Canis. Kirkness,E.P., Bafna,V., Halpern,A.L., Levy,S., Remington,K., Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and Venter,J.C.

REFERENCE Kirkness,E.P., Bafna,V., Halpern,A.L., Levy,S., Remington,K., Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and Venter,J.C.
AUTHORS Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirkness@tigr.org
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirkness@tigr.org

TITLE JOURNAL MEDLINE
PUBLMED 22875432
COMMENT

1.668
ORGANISM "Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="Taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from peripheral blood"

ORGANISM "Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="Taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from peripheral blood"

REFERENCE AUTHORS

Kim,H., Yu,Y., Wissotski,M., Yost,D., Stum,D., Rao,K., Kudrna,D., Muller,C., Soderlund,C., Bowers,J.E., Paterson,A.H. and Wing,R.

TITLE JOURNAL COMMENT

Physical mapping of the sorghum genome
Unpublished (2004)
Contact: Rod A. Wing

Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259

Email: http://genome.arizona.edu

PCR PRIMERS

FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GCG ACC CCA

Plate: 0010 Row: H column: 14
Seq primer: CAC TCA TTA GCG ACC CCA
BAC ends.

Class: BAC ends.

Location/Qualifiers
source

RESULT 37

CC732980/c
LOCUS CC732980 679 bp DNA linear GSS 23-JUN-2003
DEFINITION CCVE059rv ZM_0_71.5 KB zea mays genomic clone ZMMBMA0537U21, genomic survey sequence.

ACCESSION CC732980
VERSION CC732980.1 GI:3215333
KEYWORDS GSS
SOURCE zea mays

ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 679)

REFERENCE AUTHORS

Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

JOURNAL COMMENT

Unpublished (2002)
Other GSS: CCVE059TH
Contact: Cathy Whitelaw

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org

Seq primer: TP
Class: sheared ends.

FEATURES source
1.679
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="Taxon:45777"
/clone="ZMMBMA0537U21"
/clone_id="ZM_0_71.5_KB"
/note="Vector: pBCSK-; Site 1: HinclI; 0.7-1.5 kb methylation filtered genomic DNA library"

ORIGIN

Query Match 64.7%; Score 19.4; DB 9; Length 679;
Best Local Similarity 79.3%; Pred. No. 8.5e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TCGAGATCTTGTCAAGGCGATAGGTCTG 29
Db 585 TCCTGCATCTTCATGGATCAATAGGTTG 557

RESULT 38

CL697382
LOCUS CL697382 699 bp DNA linear GSS 13-JUL-2004

DEFINITION SP_Ba0010H14.r SP_Ba Sorghum propinquum genomic clone

SP_Ba0010H14.3', Genomic Survey Sequence.

CL697382.1 GI:50263657

ACCESSION CL697382
VERSION GSS
KEYWORDS SOURCE
ORGANISM Sorghum propinquum
Sorghum propinquum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE AUTHORS

1 (bases 1 to 699)

Kim,H., Yu,Y., Wissotski,M., Yost,D., Stum,D., Rao,K., Kudrna,D., Muller,C., Soderlund,C., Bowers,J.E., Paterson,A.H. and Wing,R.

TITLE JOURNAL COMMENT

Physical mapping of the sorghum genome
Unpublished (2004)
Contact: Rod A. Wing

Arizona Genomics Institute

University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595

Fax: 520 621 1259

Email: http://genome.arizona.edu

PCR PRIMERS

FORWARD: TAA TAC GAC TCA CTA TAG GG

BACKWARD: CAC TCA TTA GCG ACC CCA

Plate: 0010 Row: H column: 14

Seq primer: CAC TCA TTA GCG ACC CCA

BAC ends.

Class: BAC ends.

Location/Qualifiers
source

1. 699

/organism="Sorghum propinquum"
 /mol type="genomic DNA"
 /db_xref="taxon:132711"
 /clone="SP_Ba0010H4"
 /clone lib="SP_Ba"
 /note="Vector: pBelBAC1; Site_1: HindIII; Site_2:
 HindIII; Paterson lab BAC library (HindIII)"

ORIGIN

Query Match	Score 19.4;	DB 9;	Length 699;
Best Local Similarity	79.3%;	Pred. No. 8.5e+02;	
Matches	23;	Conservative	0;
	0;	Mismatches	6;
		Indels	0;
		Gaps	0;

Qy 2 CGAGGATCTGTCAAGGGAGTAGGCAGC 30
 Db 108 CGCTATCTTCAGGATCAATTGGCTGC 136

RESULT 39
 CL700732 LOCUS CL700732 704 bp DNA linear GSS 13-JUL-2004
 DEFINITION SP_Ba0061021.f SP_Ba Sorghum propinquum genomic clone
 SP_Ba0061021.s', genomic survey sequence.

ACCESSION CL700732
VERSION G1:502676707
KEYWORDS
SOURCE Sorghum propinquum
ORGANISM Sorghum propinquum; Streptophytina; Embryophyta; Tracheophyta; Bivalviflora; Viridiplantae; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
REFERENCE 1. (bases 1 to 704)
AUTHORS Kim, H., Yu, Y., Wissotski, M., Yost, D., Stum, D., Rao, K., Kudrna, D., Muller, C., Soderlund, C., Bowers, J.B., Patterson, A.H. and Wing, R.
TITLE Physical mapping of the sorghum genome
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
 Arizona Genomics Institute
 University of Arizona
 Frontiers Building Room 303, Tucson, AZ 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: http://genome.arizona.edu
PCR PRIMERS
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TAA GGC ACC CCA
PLATE: 0064
ROW: L
COLUMN: 21
SEQ PRIMER: TAA TAC GAC TCA CTA TAG GG
CLASS: BAC ends
FEATURES source
 Location/Qualifiers
 1..704
 /organism="Sorghum propinquum"
 /mol type="genomic DNA"
 /db_xref="taxon:132711"
 /clone="SP_Ba0064L21"
 /clone lib="SP_Ba"
 /note="Vector: pBelBAC1; Site_1: HindIII; Site_2:
 HindIII; Paterson lab BAC library (HindIII)"

ORIGIN

Query Match	Score 19.4;	DB 9;	Length 704;
Best Local Similarity	79.3%;	Pred. No. 8.5e+02;	
Matches	23;	Conservative	0;
	0;	Mismatches	6;
		Indels	0;
		Gaps	0;

Qy 2 CGAGGATCTGTCAAGGGAGTAGGCAGC 30
 Db 152 CGAATATCTTCAGGATCAATTGGCTGC 180

RESULT 40
 CB653308/C LOCUS CB653308 710 bp mRNA linear EST 09-APR-2003
DEFINITION OSJNEC04D10.r OSJNEC Oryza sativa (japonica cultivar-group) cDNA

clone OSJNEC04D10_3', mRNA sequence.
 CB653308.1 GI:29657033
ACCESSION
VERSION
KEYWORD SOURCE
ORGANISM Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Enhaertoidae; Orzeae; Oryza.
REFERENCE
AUTHORS Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E., Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.
TITLE Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
FEATURES source
 Location/Qualifiers
 1..710
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="Taxon:3947,"
 /clone="OSJNEC04D10"
 /tissue_type="leaf"
 /dev_stage="3 week"
 /lab_host="IDH10B"
 /clone_id="OSJNEC"
 /note="Primer: pBlueScript II KS+; Site_1: ECORI; Site_2:
 XbaI; 6 hrs after inoculation with Rice Blast (C9240-1)"
ORIGIN

Query Match	64.7%;	Score 19.4;	DB 6;	Length 710;
Best Local Similarity	79.3%;	Pred. No. 8.5e+02;		
Matches	23;	Conservative	0;	
	0;	Mismatches	6;	
		Indels	0;	
		Gaps	0;	

Qy 2 CGAGGATCTGTCAAGGGAGTAGGCAGC 30
 Db 660 CCAGGATCTCCAGGCGAGAGCTGC 632

Search completed: January 7, 2005, 12:37:44
 Job time : 1796.05 secs

